An Investigation of Attrition from Community-based Offending Behaviour Programmes

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy (or other degree as appropriate)

By

Ruth Megan Hatcher

April 2009

Acknowledgements

My thanks go out to my supervisor, Professor James McGuire, for his patient support over the years. Despite many obstacles along the way, James has been positive and reassuring at times when I felt that this thesis would never be completed. I have no doubt that if it wasn't for James' encouragement, I would have given up the notion of this thesis a long time ago.

I am also grateful to the University of Leicester for granting me a period of study leave in September 2007 which permitted time for the data collection which forms part two of this thesis. My colleagues at the University over the last six years have also been hugely encouraging to me and their help in supporting me to the completion of this work really has been invaluable.

The local Probation Service (which I cannot name due to confidentiality reasons) has also played a large role in this thesis. I spent four months based within the programme team's office during the winter of 2007. All of the staff, in particular the Programme Manager and the administrative staff, were welcoming and exceptionally helpful and I thank them for that. I am also grateful to the eleven individuals who gave up their time to speak with me about their experiences of the Enhanced Thinking Skills programme. Whilst their contributions have not been presented within this final volume, I will ensure that the time that they gave to this work is not wasted by writing up their contributions within subsequent papers.

I also owe a massive debt of thanks to my close friends and family (and especially my boyfriend James) for putting up with me in recent months (or perhaps years!). I apologise for not being there when I should have been and, for when I was there, talking endlessly about the endless PhD! I am extremely thankful for your patience and support.

Finally, I dedicate this thesis to my Mum and Dad for inspiring me to do the best I can.

An Investigation of Attrition from Community-based Offending Behaviour Programmes

Ruth Megan Hatcher

Abstract

Objectives: This thesis investigates attrition from community-based cognitive-skills offending behaviour programmes. Part one of the thesis investigates the influence of attrition on the reconviction outcomes of those sentenced to either the Enhanced Thinking Skills (ETS) programme or the Think First programme. The data are investigated to determine whether programme dropout is detrimental to reconviction outcomes. Part two undertakes a more thorough analysis of the characteristics of programme completers, non-completers, and non-starters of the ETS programme. The three groups are compared to assess for differences in relation to demographic, psychometric, criminal history, and offender need variables. The role of organisational, or process, factors in attrition is also investigated. Finally, the reasons recorded in probation files for non-attendance at the ETS programme are examined.

Methods: Part one comprises two chapters and utilises a quasi-experimental design. Data relating to a national sample of offenders sentenced to an offending behaviour programme and a comparison group of offenders sentenced to probation but not required to undertake a programme and matched on a one-to-one basis to the experimental group are utilised. Part two utilises data relating to offenders sentenced to the ETS programme within one probation area. The focus on one locality permitted a rich analysis of the factors associated with attrition. Analyses undertaken include tests of association (correlations, chi-square), parametric and non-parametric tests for differences (t-tests, ANOVAs, Kruskal Wallis, Mann Whitney), logistic regression (binary and multinomial), and calculations of effect sizes.

Results: The analyses in part one provided tentative evidence of a negative impact of attrition on reconviction outcomes. This trend remained after controlling for those variables found to significantly differ between groups. The analyses within part two indicated that programme non-starters were more criminogenic than programme completers and non-completers. The non-completers, however, were the youngest of the groups and hence had less of a history but displayed a similar rate of offending as the non-starters. In relation to process factors, there was a significant association between appropriateness of targeting and attrition; offenders with risk of reconviction scores above the recommended criteria were most likely to dropout. Finally, a third of dropouts could not, should not, or were not able to attend due to the unavailability of a programme place, a further third were already in breach of their order or had committed a further offence, and the final third could and should attend, were not apparently in breach but still failed to commence.

Conclusions: Programme dropouts produce worse reconviction outcomes than programme completers and matched comparisons. In evaluating correlates of attrition, dropouts are more criminogenic than programme completers. However, process factors were also associated with programme attrition. Research should investigate the impact of individual and process factors on attrition further and should use these findings to inform the debate concerning the influence of programme attrition (and hence completion) on reconviction outcomes. It is anticipated that the findings will inform clinical practice and the treatment readiness and intervention outcome research literatures.

Contents

Acknowledgments	i
Abstract	ii
List of tables and figures	vi
Chapter One: Literature Review	1
PART ONE: Attrition: Investigation of attrition and reconviction outcomes utilising national probation data	40
An Introduction to Part One	41
Chapter Two: Attrition: An analysis of reconviction outcomes	42
Introduction	43
Hypotheses	47
Method	48
Results	54
Discussion	67
Chapter Three: Attrition: A Comparison of completers, non-completers, and non-starters	73
Introduction	74
Hypotheses	81
Method	82
Results	87
Discussion	101

PART TWO: Investigation of attrition utilising data from one probation area 108

An Introduction to Part Two 109

Chapter Four: An investigation of offender demographic and psychometric variables and their relationship with attrition 110

- Introduction 111
- Hypotheses 115
- Method 116
- Results 122
- Discussion 129

Chapter Five: An investigation of offence-related variables, criminogenic and non-criminogenic need and their relationship with attrition 135

- Introduction 136
- Hypotheses 138
- Method 139
- Results 142
- Discussion 189

Chapter Six: An investigation of process related variables and their relationship with attrition 197

- Introduction 198
- Hypotheses 203
- Method 204
- Results 206
- Discussion 230

Chapter Seven: An analysis of the reasons for attrition 237

Introduction 238

Method 242

Results 244

Discussion 272

Chapter Eight: General Discussion 278

References 297

Appendices 319

Word count: 94 398

List of Tables and Figures

Tables:

Table 2.1. Descriptive Statistics by Treatment Allocation Group 52
Table 2.2: Logistic regression of reconviction as a function of time at risk and treatment group (experimental, comparison) 55
Table 2.3. Descriptive Statistics by Treatment Received Groups 56
Table 2.4: Logistic regression of reconviction as a function of time at risk and treatment group (comparison, completer, non-completer, non-starter) 57
Table 2.5: Odds ratio for each pairwise comparison 58
Table 2.6: Logistic regression of reconviction as a function of time at risk and treatment group (completer, matched comparison) 59
Table 2.7: Logistic regression of reconviction as a function of time at risk and treatment group (non-completer, matched comparison) 60
Table 2.8: Logistic regression of reconviction as a function of time at risk and treatment group (non-completer, matched comparison) 61
Table 2.9: Descriptive Statistics by Treatment by Starter and Non-starter groups 61
Table 2.10: Reconviction by group 62
Table 2.11: Logistic regression of reconviction as a function of OGRS2, time at risk, and treatment group (starters, non-starters) 63
Table 2.12: Descriptive Statistics by Treatment by Completer and Dropout groups 64
Table 2.13: Reconviction by group 64
Table 2.14: Logistic regression of reconviction as a function of age, OGRS2, time at risk, and treatment group (completers, dropouts) 65
Table 3.1. Descriptive Statistics by Treatment Groups 87
Table 3.2. OGRS2 criteria classification by group 88
Table 3.3. OGRS2 appropriateness by group 89
Table 3.4. Motivation to complete the programme mean scores by group. 89
Table 3.5. School leaving age by group 90
Table 3.6. Employment status by group 90

Table 3.7. Accommodation status by group 91
Table 3.8. Literacy problems by group 91
Table 3.9. Need scores by group 92
Table 3.10. Correlations between age, OGRS2, number of previous convictions and need scores. 93
Table 3.11. Mean suitability scores by suitability group 93
Table 3.12. Suitability by group (three groups). 94
Table 3.13. Mean suitability score by group (three group) 95
Table 3.14. Multinomial logistic regression of programme completion as a function of OGRS2 96
Table 3.15. Multinomial logistic regression of programme completion as a function of OGRS2 appropriateness 98
Table 3.16: Logistic regression of reconviction as a function of time at risk, OGRS2, and treatment group (comparison, completer, non-completer, non-starter) 100
Table 3.17: Logistic regression of reconviction as a function of time at risk, OGRS2 appropriateness classification, and treatment group (comparison, completer, non-completer, non-starter) 101
Table 4.1: Descriptive Statistics by Group 122
Table 4.2: Gender by Group 124
Table 4.3: Ethnicity by Group 124
Table 4.4: Pre-programme psychometric scores by group 126
Table 4.5: Multinomial logistic regression of programme completion as a function of Age, OGRS2, Racial Intimacy and Aggressive 128
Table 5.1: Distribution of offence categories for index offences across the Sample 142
Table 5.2: Distribution of offence categories of index offences across groups (%) 143
Table 5.3: Distribution of offence categories of index offences (with the collapsed category of acquisitive crime) across groups (%) 144
Table 5.4: Breach by group 145
Table 5.5: Mean number of court appearances at which convicted by group 146
Table 5.6: Mean age of first conviction by group 147

Table 5.7: Mean Copas Rate by group 148
Table 5.8: Mean number of custodial sentences by group 148
Table 5.9: Mean number of offence categories for which convicted by group 149
Table 5.10: Mean number of previous violent and sexual convictions by group 150
Table 5.11: Mean OASys total score by group 151
Table 5.12: Mean OASys criminogenic and non-criminogenic needs scores by group 151
Table 5.13: OASys individual need scores by group 153
Table 5.14: Kolmogorov-Smirnov output for the OASys individual need scores 154
Table 5.15: OASys individual criminogenic need scores by group 156
Table 5.16: Kolmogorov-Smirnov output for the OASys individual criminogenic need scores 157
Table 5.17: OASys individual non-criminogenic need scores by group 159
Table 5.18: Kolmogorov-Smirnov output for the OASys individual non-criminogenic need scores 160
Table 5.19: Distribution of OASys risk classifications by group 161
Table 5.20: Multinomial logistic regression of programme completion as a function of previous breach, OASys total score, number of court appearances at which convicted under 18 and age of first conviction. 163
Table 5.21: Multinomial logistic regression of programme completion as a function of previous breach, OASys criminogenic score, number of court appearances at which convicted under 18 and age of first conviction. 165
Table 5.22: Multinomial logistic regression of programme completion as a function of previous breach, OASys total score, and Copas Rate 167
Table 5.23: Distribution of offence categories of index offences across starter and non-starter groups (%) 168
Table 5.24: Distribution of offence categories of index offences (with the collapsed category of acquisitive crime) across starter and non-starter groups (%) 169
Table 5.25: Breach by starter and non-starter group 170
Table 5.26: Mean number of convictions by group 170
Table 5.27: Mean age of first conviction by starter and non-starter group 171

Table 5.28: Mean Copas Rate by group. 171
Table 5.29: Mean number of custodial sentences by group 172
Table 5.30: Mean number of offence categories for which convicted by starter and non-starter group 172
Table 5.31: Mean number of previous violent and sexual convictions by starter and non-starter group 173
Table 5.32: Mean OASys total score by starter and non-starter group 174
Table 5.33: Mean OASys criminogenic and non-criminogenic needs scores by starter and non-starter groups 175
Table 5.34: OASys individual need scores by starter and non-starter group 176
Table 5.35: Kolmogorov-Smirnov output for the OASys individual need scores 177
Table 5.36: Kolmogorov-Smirnov output for the OASys individual criminogenic need scores 178
Table 5.37: OASys individual criminogenic need scores by starter and non-starter group 179
Table 5.38: Kolmogorov-Smirnov output for the OASys individual non-criminogenic need scores 180
Table 5.39: OASys individual non-criminogenic need scores by starter and non-starter groups 181
Table 5.40: Distribution of OASys risk classifications by starter and non-starter groups 182
Table 5.41: Logistic regression of programme starting as a function of previous breach, number of court appearances at which convicted under 18, age of first conviction, and OASys total score. 184
Table 5.42: Logistic regression of programme starting as a function of previous breach, number of court appearances at which convicted under 18, age of first conviction, and OASys criminogenic score 185
Table 5.43: Logistic regression of programme starting as a function of previous breach, Copas Rate, and OASys criminogenic score. 186
Table 5.44: Summary of significant findings 187
Table 6.1: Order type by group 206
Table 6.2: Mean number of attempts by group 207

Table 6.3: Cumulative completion status outcomes in relation to the number of attempts 208
Table 6.4: OGRS2 criteria classification by group 209
Table 6.5: OGRS2 appropriateness by group 209
Table 6.6: Mean time from order to programme start by group. 210
Table 6.7: Distribution of group by time of programme. 211
Table 6.8: Distribution of group numbers of programme sessions per week 212
Table 6.9: Mean Level of Engagement and Understanding scores by group 212
Table 6.10: Multinomial logistic regression of programme completion as a function of number of previous attempts on this order and OGRS2 classification (too low, appropriate, too high) 215
Table 6.11: Logistic regression of group (completer, non-completer) as a function of OGRS2 classification (too low, appropriate, too high), number of attempts on the order, and average level of engagement scores (completers, dropouts) 217
Table 6.12: Age by completion and appropriateness groups 218
Table 6.13: OGRS2 by completion and appropriateness groups 219
Table 6.14: Pre-programme psychometrics by completion and appropriateness groups 221
Table 6.15: Copas rate by completion and appropriateness groups 223
Table 6.16: Total OASys score by completion and appropriateness groups 224
Table 6.17: Total OASys criminogenic score re by completion and appropriateness groups 227
Table 6.18: Total OASys non-criminogenic score re by completion and appropriateness groups 230
Table 7.1: Reasons for type one non-attendance 244
Table 7.2: Reasons for type two non-attendance 245
Table 7.3: Reasons for type three non-attendance 247
Table 7.4: Reasons for Type One non-attendance 249
Table 7.5: Reasons for Type Two non-attendance 250
Table 7.6: Reasons for Type three non-attendance 251

Table 7.7: Age by Attrition type (type one, type two, type three) and Group (non-starter, non-completer) 253
Table 7.8: OGRS2 by Attrition type (type one, type two, type three) and Group (non-starter, non-completer) 254
Table 7.9: Pre-programme Racial Intimacy by Attrition type (type one, type two, type three) and Group (non-starter, non-completer) 255
Table 7.10: Pre-programme Aggressiveness by Attrition type (type one, type two, type three) and Group (non-starter, non-completer) 256
Table 7.11: Previous breach by attrition (type one, type two, type three) group 256
Table 7.12: Previous breach by attrition (type one, type two, type three) and dropout group (non-starter, non-completer) 257
Table 7.13: Total number of previous convictions and number of previous convictions prior to age 18 by attrition (type one, type two, type three) and dropout group (non-starter, non-completer) 258
Table 7.14: Age at first conviction by attrition (type one, type two, type three) and dropout group (non-starter, non-completer) 259
Table 7.15: Copas rate by Attrition type (type one, type two, type three) and Group (non-starter, non-completer) 260
Table 7.16: OASys total score by Attrition type (type one, type two, type three) and Group (non-starter, non-completer) 261
Table 7.17: OASys risk classification by attrition (type one, type two, type three) group 261
Table 7.18: OASys risk classification by attrition (type one, type two, type three) and dropout group (non-starter, non-completer) 262
Table 7.19: OASys criminogenic need score by Attrition type (type one, type two, type three) and Group (non-starter, non-completer) 263
Table 7.20: Mean number of attempts at ETS on this order by Attrition type (type one, type two, type three) and Group (non-starter, non-completer) 264
Table 7.21: OGRS2 appropriateness by attrition (type one, type two, type three) group 264
Table 7.22: OGRS2 appropriateness by attrition (type one, type two, type three) and dropout group (non-starter, non-completer) 265

Table 7.23: Age and OGRS2 by Attrition type (3A, 3B, 3C) amongst the non-starter group 266
Table 7.24: Pre-programme Racial Intimacy and Aggressiveness by Attrition type (3A, 3B, 3C) amongst the non-starter group 267
Table 7.25: Previous breach by attrition group (3A, 3B, 3C) amongst the non-starter group 267
Table 7.26: Total number of previous convictions and number of previous convictions prior to age 18 by attrition group (3A, 3B, 3C) amongst the non-starter group 268
Table 7.27: Age at first conviction, Copas rate, OASys total score and OASys criminogenic score by attrition group (3A, 3B, 3C) amongst the non-starter group 269
Table 7.28: OASys risk classification by attrition group (3A, 3B, 3C) amongst the non-starter group 270
Table 7.29: Mean number of attempts at ETS on this order by attrition group (3A, 3B, 3C) amongst the non-starter group 271
Table 7.30: OGRS2 appropriateness by attrition group (3A, 3B, 3C) amongst the non-starter group 271

Figures:

Figure 6.1: Scatterplots of OGRS2 risk score against Total OASys score by completion group (too high group only) 225
Figure 6.2: Scatterplots of OGRS2 risk score against Total OASys score by completion group (appropriate group only) 225
Figure 6.3: Scatterplots of OGRS2 risk score against OASys criminogenic need score by completion group (too high group only) 228
Figure 6.4: Scatterplots of OGRS2 risk score against OASys criminogenic need score by completion group (appropriate group only) 229

Chapter One

Literature Review

Literature Review

History and development of the National Probation Service

The role of the modern National Probation Service (NPS) for England and Wales is to "protect the public, reduce re-offending, [ensure] proper punishment of offenders within the community, ensure offenders' awareness of the effects of crime on the victims of crime and the public, and the rehabilitation of offenders" (National Probation Service, 2003a, p. 1). The NPS attempts these aims, under the umbrella of the National Offender Management Service, through the administration of a number of community-based initiatives which are sanctioned by the criminal justice and court system.

Although probation as a form of provisional release from prison was introduced as long ago as 1907, its designation as a sentence of the court was not formalised until the Criminal Justice Act of 1991. However, the development of a large proportion of present day community sentences can be traced to the new Labour government of 1997 and the implementation of the Effective Practice Initiative (Hedderman, 2007) which was predicated on the influential Underdown (1998) report for the Probation Inspectorate. Underdown's survey of 267 local and 'effective' probation initiatives alarmingly found only four of these to have been properly evaluated and to be producing positive outcomes (Raynor, 2003; Raynor & Vanstone, 2007). Such a disappointing picture of the value of the probation service ultimately led to the development of a centrally managed service. In 2001, the 54 separate Probation Services were amalgamated into the NPS which was centrally governed by the National Probation Directorate (Hedderman, 2007). This new government agency was focussed on the introduction and development of initiatives to protect the public and reduce crime that were based within the available effectiveness literature (Raynor & Vanstone, 2007).

In the meantime, a suite of 'Pathfinder' programmes had been under development and had been aided by funding from the government's Crime Reduction Programme (CRP). The CRP was "the most ambitious, best-resourced and most comprehensive effort for driving down crime ever attempted in a Western developed country" (Homel, Nutley, Webb, & Tilley, 2004, p. i) and focussed on five themes:

"...preventing young people becoming offenders of the future, tackling crimes in communities..., developing products and systems that are resistant to crime, more effective sentencing practices, working with offenders to ensure that they do not reoffend" (Homel et al., 2004, p. v). Under the auspices of this final theme, the NPS 'Pathfinder' programmes were identified and piloted. The intention was to take principles from the "What Works" literature and apply them within practice (Hedderman, 2007) to develop "constructive approaches to working with offenders to prevent crime" (Hollin & Palmer, 2006, p. xi). As such, projects focussed on four key areas were established: improving offenders' basic skills, the resettlement of short-term prisoners on release, pro-social community service supervision, and the delivery of offending behaviour interventions.

Offending behaviour programmes within the NPS

The fourth of these 'Pathfinders' focussed on the provision of evidence-based offending behaviour programmes. The aim of these programmes was to reduce offending behaviour through the use of rehabilitative efforts. Defined as "structured approach[es] to helping offenders to acquire the skills and knowledge which can help them to stay out of trouble" (Canton & Hancock, 2007), the development of offending behaviour programmes was based within cognitive social learning theory (Bandura, 1977). Social learning theory stipulates that human beings learn not only from direct experience but also indirectly from the observation of the behavioural outcomes of others (McGuire, 2006). As such, deviant behaviour, such as delinquency and criminal behaviour, rather than being resultant of psychopathology, is acquired and maintained through socialisation and reinforcement in much the same way as other behaviour is learned (Akers, Krohn, Lanza-Kaduce & Radosevich, 1979). Given this premise, there comes the possibility that new skills and new forms of pro-social behaviour can be learned by criminally deviant individuals to provide them with the 'toolkit' to avoid future criminal behaviour.

Thus social learning concepts were combined with the findings of cognitive and developmental research (McGuire, 2006): offenders were found to have lower levels of interpersonal, self management, problem solving, and social skills than the general population (Ross & Fabiano, 1985; Spivack & Levine, 1963). It is hence upon this

theoretical and evidence base that present day offending behaviour programmes were formed. Offending behaviour programmes aim to equip offenders with interpersonal, cognitive and self-management skills to help them avoid future deviant and criminal behaviour. For example, the Think First programme, a general offending behaviour programme (as opposed to one designed for a specific type of offence, such as sex offending, or drink driving), designed by McGuire (2005), focuses on the provision of social problem solving skills, such as problem awareness, alternative-solution thinking, consequential thinking, and perspective taking. The components of self management and social interaction training, alongside values education comprise the remainder of the programme. It is the acquisition of problem solving skills, however, which is deemed to be key and the basis upon which the self management, social interaction, and values education training sits (Home Office, 2000a).

Programmes such as Think First were developed on the basis of the findings of statistical reviews of large numbers of evaluations of such interventions. The meta-analyses of juvenile and adult correctional treatment published towards the end of the twentieth century (see for example, Andrews, Zinger, Hoge, Bonta, Gendreau & Cullen, 1990; Antonowicz & Ross, 1994; Dowden & Andrews, 1999a, 1999b, 2000; Izzo & Ross, 1990; Lipsey, 1992, 1999; Redondo, Sanchez-Meca & Garrido, 1999; Whitehead & Lab, 1989) had highlighted that programmes for offenders could indeed reduce recidivistic outcomes if they are designed and implemented in such a way as to facilitate this. Given the similarity in the findings of the meta-analyses, attempts to synthesise the outcomes into a set of evidence based principles upon which effective interventions could be based followed. It is upon these 'principles of effective practice' (Hollin & Palmer, 2006) or 'principles of human service' (Andrews, 1995, 2001) that the accreditation criteria for programmes within the prison and probation services of England and Wales were founded. These ten criteria are outlined by Lipton, Thornton, McGuire, Porporino and Hollin (2000) as:

1. The programme should be based within an explicit, empirically based model of change;
2. The programmes should target those needs which have been shown by research to be linked to offending behaviour (criminogenic needs);

3. The programmes should be motivational and be responsive to the needs of the offenders targeted by the programme;
4. The methods used to target the criminogenic needs should be supported by empirical evidence which verifies their effectiveness with the target population;
5. The programme should be skills oriented and should provide offenders with those skills that support a pro-social life;
6. The programmes should address a range of conceptually different treatment targets;
7. The sequencing, intensity, and dosage of the programme should relate to the risk and seriousness of the target group of offenders;
8. The programmes should be embedded within a system that supports the rehabilitative effort;
9. There should be ongoing and built-in monitoring of the programme;
10. There should be ongoing and built-in evaluation of the programme.

The goal of accreditation, which was initially overseen by the General Accreditation Panel (which later became known as the Joint Accreditation Panel and later still the Correctional Services Accreditation Panel), was therefore to ensure that offender programming was based within evidence based guidelines and that those that had not been substantiated as effective models of intervention were discouraged (Lipton et al., 2000; Hollin & Palmer, 2006). The aim of the 'Pathfinder' offending behaviour programme pilots was therefore to inform the implementation and management, and to permit thorough evaluation, of such programmes within the community of England and Wales. The findings of the evaluations that ran alongside the pilots (e.g. Hollin, Palmer, McGuire, Hounsome, Hatcher, Bilby, & Clark, 2004) would provide evidence for the Joint Accreditation Panel.

As McGuire writes, however, "In its speed and scale, the dissemination of structured offending behaviour programmes within criminal justice settings has been little short of remarkable" (McGuire, 2006, p. 69). Raynor and Vanstone (2007) agree that the pace and scale was unprecedented. Speaking of the Pathfinder projects in general, Raynor and Vanstone stated that "No correctional service anywhere in the world had tried to implement 'What Works' principles on such a scale, at such a speed

and subject to such comprehensive scrutiny and evaluation" (p. 72). As such, the pilot offending behaviour programmes were soon being delivered on a national scale. The suite of available programmes encompassed general offending behaviour programmes, such as the Think First programme above, and specialised programmes for those who had committed specific types of offences, such as sex offences, aggression related offences, and substance related offending. Four general offending behaviour programmes were available: three group-based programmes, Think First, Reasoning and Rehabilitation (R&R) and Enhanced Thinking Skills (ETS), and one, the Priestley One-to-One programme, designed to be delivered on an individual basis.

The R&R programme was devised in Canada by Ross and Fabiano (Ross & Fabiano, 1985; Ross, Fabiano, & Ewles, 1988). The programme comprises 38 two hour sessions at a rate of two to four sessions a week (McGuire, 2006) and focuses on interpersonal cognitive problem-solving skills, social skills, emotional management, creative thinking, critical reasoning, value enhancement and metacognition (McGuire, 2006). A shorter version of this programme was developed within England and Wales by prison staff to meet the need of lower risk offenders. This 20 session programme was the ETS programme (Clark, 2000). Both programmes were developed, alongside the Think First programme, to meet the Correctional Services Accreditation Panel criteria for community based interventions for offenders. Each probation area was encouraged to select for delivery one of the general offending behaviour programmes (Think First, Reasoning and Rehabilitation, Enhanced Thinking Skills, or Priestley One to One) and to consider implementing the available offence specific programmes if they met the need profiles of offenders within their locality.

Attrition: Facts and figures

Whilst the delivery of accredited programmes grew apace in England and Wales, the rapid implementation of such an initiative was not without its problems (Hollin, McGuire, Palmer, Bilby, Hatcher, & Holmes, 2002a; Hollin, McGuire, Palmer, Bilby, Hatcher, & Holmes, 2002b; Raynor, 2004, 2008). The process evaluation of the early pilot Pathfinder programmes found referrals were often inappropriate, resources were stretched, attention to programme integrity was lacking, and the accommodation and administrative support for programmes were poor (Hollin, et al., 2002a, 2002b). Perhaps

allied to these issues, the research also concluded that the programme attrition, or non-completion, rate of accredited programmes was particularly high: for example, the offence specific programmes returned non-completion rates ranging from 25% to 80% (Hollin, et al, 2002a).

Attrition is neither a new problem nor one that is specific to offending behaviour programmes, the UK, or community settings. An early review of medical and psychiatric treatments (Baekeland & Lundwall, 1975) found attrition rates of between 26% and 68% for community based services and between 23% and 39% for inpatient services. Likewise, a review of the early substance abuse treatment literature (Stark, 1992), found dropout rates generally over 50% within the first month of treatment. Within North America, offending behaviour programmes have produced non-completion rates of 37.6% within the secure estate (Wormith & Olver, 2002) and 40% amongst parolees (van Voorhis, Spruance, Ritchy, Listwan, & Seabrook, 2004). Likewise, prison based research relating to a substance use programme delivered across 20 US prisons reported average dropout rates at 16% but with a range of 3% to 44% across the prison estate (Pelissier, Camp, & Motivans, 2003). Within the UK, Cann, Falshaw, Nugent & Friendship (2003) reported a total rate of cognitive skills programme non-completion of just 11% across their prison sample. These latter findings highlight that completion rates for offenders on interventions within prisons tend to be more favourable than those within community settings.

Despite such knowledge, the magnitude of the attrition rates reported by Hollin et al. (2002a) for court mandated community based programmes was alarming; however, even more so was the likelihood that these figures in all probability underestimated the extent of the phenomenon. Clearly, if offenders on probation are not present to take part in an intervention programme, it cannot achieve its intended effects. Several questions have therefore been raised concerning attrition, focusing principally on what may be the causes of it and how it might be reduced. This thesis addresses those questions and reports on a series of research studies designed to investigate factors that may influence attrition. In this first chapter, existing literature relevant to this area will be reviewed and critically examined.

The figures produced by Hollin et al. were based on practitioner estimates of *within*-programme dropout: that is the count of those offenders who commenced a

programme but failed to complete it (to be referred to within this thesis as 'programme non-completers'). As such, those offenders who failed to attend any portion of the intervention (to be referred to within this thesis as 'programme non-starters') were not considered within these data. A 2002 national evaluation, however, reported attrition rates combining both non-starters and non-completers to be as high as two thirds of those sentenced to attend (Hollin, et al., 2004). The Think First and ETS programmes were reported as having the similar attrition rates (66% and 69% respectively) with the R&R programme faring worse: only one in five of those sentenced to the R&R programme completed it (Palmer, McGuire, Hounsome, Hatcher, Bilby, & Hollin, 2007). By 2003-4, it was possible to provide a breakdown of attrition at the two distinct stages: of those receiving general offending behaviour programme court orders, 48.5% failed to commence the programme, 23.3% commenced but did not complete, and a meagre 28.2% completed the programme to which they were sentenced to by the courts (Hollin, McGuire, Hounsome, Hatcher, Bilby, & Palmer, 2008).

These national surveys of programme attrition can be supplemented with the findings of a series of local studies that examined the issue of attrition with a view to informing local practices. For example, Hazeltine, Walker, Nickells, and Gregg (2003) examined attrition from the Think First programme within the Bedfordshire Probation Area. Of offenders referred to attend the Think First programme, just 27% completed the programme. Of the remainder almost one quarter never received an instruction to attend a programme and another quarter dropped out prior to the group sessions (these two statistics combined takes the non-starter rate to 52%). An additional 21% dropped out during the group sessions (non-completers). Given additional complexities in the data (for example, these figures relate to referrals rather than to offenders: some offenders would have received more than one referral) such statistics should be taken as tentative indications of the extent of dropout.

Sussex Probation Area (2003) reported an overall attrition rate from their accredited programmes of one in two. However, definitional ambiguity within this report does not allow for specificity; attrition is presented as the percentage of "dropouts" to "starters" but neither of these categories has been defined. What is clear, however, is that, in using starters as the reference category, the attrition statistic would

not include non-starters of programmes. As such, the figure of 51% underestimates the true attrition rate within that Probation Area at that time.

A further report commissioned by Northumbria Probation Area in 2001 (Westmarland, Hester, Reid, Coulson, & Hughes, n.d.) reported that pre-programme Think First attrition rates (i.e. non-starters only) were as high as 58%; greater than the national figures but similar to those reported by Hazeltine et al (2003). By 2002, however, the overall attrition rate, this time incorporating non-starters and non-completers was down to 37%: 16% non-starters and 21% non-completers. This represents a reduction in the non-starter rate between the two measurement points of a commendable 42%; unfortunately no information was provided by Westmarland et al. as to how such a reduction in attrition was achieved. Additionally, it is not clear from their report how the treatment group was defined; it is unclear whether the counting started when offenders received an order to attend the programme or when they were referred to a particular delivery run of programme.

It would seem therefore that whilst informative, the findings of these local reports of programme attrition are limited because they each fail to define attrition or their attrition groups comprehensively. Where these concepts are defined, they differ across probation areas such that meaningful comparisons are not possible. Despite these ambiguities, it is clear that attrition from offending behaviour programmes, at that time, was substantial and concerning for those charged with managing these programmes both locally and centrally.

Such concern was reflected in the commissioning of a National Offender Management System (NOMS) sponsored evaluation of programmes using data from the Interim Accredited Programme System (IAPS: Hollis, 2007). IAPS is a national computer database which holds locally inputted information relating to programme management and throughput and it was used to compile national attrition and completion rates of the general offending behaviour programmes (including the Cognitive Skills Booster programme[1], the Priestley One to One programme, ETS, Think First and R&R) from a sample of almost 13000 offenders. Hollis reported a much reduced non-starter rate of

[1] The Cognitive Skills Booster programmes was designed for offenders who had completed one of the general offending behaviour programmes and "enables the offender to refresh and apply the skills learned...to real life situations" (Dawson, Walmsley, & Debidin, 2005, p iii.)

just less than one in five (compared to one in two in 2001/02 as reported by Hollin et al. (2008)). However, these improved commencement rates did not result in higher completion rates: these remained relatively static at 28.7% (Hollin et al. reported completion rates of 28.2%). From the work of Hollis therefore, the overall rate of attrition had not altered but the nature of it had: whilst pre-programme attrition had reduced substantially, 52.5% of offenders now started but did not complete their programme.

A few words of caution should be expressed about Hollis's (2007) research. One of the advantages of the research is the size of the sample utilised; the irony of this, however, is that it produces its own difficulties. With such vast quantities of data, it becomes impractical to check, on an individual level, the accuracy of the data. The author herself admits that the classifications of offenders as completers, non-completers and non-starters on the basis of IAPS codings may have resulted in incorrect classifications. In addition, nearly ten percent of the sample was removed from the analysis because ambiguities in the data rendered it impossible to distinguish whether they were non-starters, non-completers or completers. If it is assumed that these are likely to fall randomly into these three groups then this exclusion might not significantly impact on the results. However, this is a large assumption to make in a service where there is pressure to record every completer accurately to reach ambitious funding-related targets (Raynor, 2004). It is perhaps more likely therefore that these individuals would have fallen into the non-starter or non-completer groups.

What is clear, however, is that the problem of attrition from offending behaviour programmes within the community services of England and Wales is sizeable and requires attention. National and local reviews of the problem have highlighted issues with programme non-start and within-programme dropout. As such, these figures indicated that research is needed to investigate this issue, its causes, and the potential impact of it on programme outcomes.

Attrition: Definitional concerns

Nunes and Cortoni (2006a) argue that the research within the field of attrition from offending behaviour programmes has been thwarted by definitional issues. As argued above in respect of the local probation studies of attrition, Nunes and Cortoni

submit that the utilisation of different definitions of attrition makes it difficult, if not impossible, to compare research findings to further understand the concept. Not only have definitions varied between studies but the heterogeneity of programme dropouts has often been ignored within research (e.g. Craissati & Beech, 2001; Geer, Becker, Gray, & Kraus, 2001; van Voorhis et al., 2004). As such, research has often amalgamated the programme non-starters and non-completers into one group of 'dropouts' (e.g. Hollin et al., 2004) or have failed to consider the non-starters entirely (e.g. Sussex Probation Service, 2003).

Where research has attempted to separate out different types of programme dropout, qualitative differences have been found between such groups. For example, Pelissier, Camp and Motivans (2003) conducted a multisite evaluation of treatment non-completion within 20 prisons in the USA to determine whether individual and process variables distinguished completers from non-completers of substance use programmes. The authors were able to predict not only non-completion but also the different types of non-completion, those who had been removed from treatment and those who had withdrawn, from individual and process related variables. These differences, they argue, would have "been masked if we had not differentiated between the two types of non-completion" (p. 139).

Within UK community based research, and in an effort to understand the concept better, Stephens (2003) has proposed an action-oriented typology of programme attrition to inform practice and research. The typology not only separates non-starters from non-completers but also classifies pre-programme attrition (i.e. non-starters) into a further three types; those who *could not* attend treatment, those who *could* attend but *should not* attend, and those who *could and should* but *did not* attend. Those offenders in the first group, those that could not attend treatment, will have received the order to attend a programme but circumstances would subsequently have not allowed participation. 'Type one' factors, as Stephens calls them, renders programme attendance not possible. Such factors may include transference out of the area, hospitalisation or death, subsequent custody for an alternative offence, or the unknown whereabouts of the offender.

Type two attrition, Stephens argues, relates to those offenders that had the ability to attend the programme (they had no Type one factors present) but should not

attend due to unsuitability for programme work. After the court has sentenced an offender to attend a programme, it is possible that further assessment may indicate unsuitability, or a change in circumstances may preclude an offender attending the programme (National Management Manual, 2000). For example, evidence of low intellectual ability, mental health problems, or substance use issues may come to light and indicate the offender may not be able to cope in a group setting. Thus attrition at this stage may not be the personal choice of the offender but instead an institutional determination which rightly complies with the targeting guidelines for the programme.

Type three offenders, Stephens argues, are those which represent 'true' attrition from programmes. These offenders are both able and considered suitable to attend a programme but for whatever reason have not attended it. Stephens proposes that there may be a "complex and inter-related set of factors" (p. 236) which contributes to true attrition and that these factors may be more difficult to determine than type one or type two factors. Prior to conception of this model Stephens had not researched type three factors. However, despite this she proposes there to be three distinct sub-categories. The first of these relates to programme design and proposes that aspects of the programme (such as the requirement to complete pre-programme psychometrics) or the session contents may "present an obstacle for some offenders" (p. 238). Second, implementation or organisational issues may contribute to treatment non-compliance. Stephens argues that issues such the lack of availability of a group due to staff shortages, or clashes of groups with employment or child care commitments, are impediments to treatment attendance which should be addressed by the organisation. Stephens highlights offender motivational factors as the third and final type 3 factor and proposes the solution of "more and/or better motivational work" (p. 239).

Whilst the 'true' attrition, or type three factors, within Stephen's model may not be evidence based and hence require further consideration, the logical categorisation of non-starters of programmes allows for the sifting of offenders not able to attend or not suitable for the programme from those who were available, eligible, and suitable but chose not to attend. Such structure allows the consideration of distinct contributors to attrition and provides a useful framework for future research within this field.

Stephens and Turner (2004) have built on this conceptual work by conducting research utilising the model above. Of those individuals who had not commenced a

programme within three months of the start of their court order, 11% were classified as type one offenders (not able to attend) and nine percent as type two (able to attend but shouldn't). Type three offenders were classified into three groups. The first of these were those for whom a group was not available either due to insufficient places on the programme or because those places that were available were not appropriate to the offender's circumstances: for example, the programme clashed with the offenders' employment or they were waiting for a female or black/Asian group. This group comprised 36% of the non-starters. The second Type three group comprised those for whom a group place was available but they were in breach of their order (29%) and the third Type three group were those for whom a group place was available, were not in breach but still did not attend session one (16%). It would seem that from the data available to the researchers it was not possible to categorise the Type three reasons into the original categories suggested by Stephens (2003; programme issues, implementation or organisational issues, and offender motivation issues). Such classification would necessitate speaking with each offender which is both time consuming and, in cases where the offender is non-contactable or incarcerated, impractical. Notwithstanding these issues, it would be useful to undertake further research utilising a classification system such as that proposed by Stephens (2003) to determine whether subsets of programme dropouts exist and whether such subsets result from different causal mechanisms.

What the above research indicates is that not all attrition is traceable to the offender. Indeed, from Stephens and Turner's (2004) figures it can be established that only approximately 45% of pre-programme attrition was initiated by the offender (those for whom a place was available but were in breach plus those who failed to attend: 29% and 16% respectively). When considered in the wider context, this information starts to clear some of the fog surrounding attrition and the causes of it; if approximately half of those sentenced to a programme fail to commence it (Hollin et al., 2008; Westmarland et al., n.d.) and over half of these non-starts are the direct result of organisational issues, the argument that programmes completers are those that "would do well anyway" (Debidin & Lovbakke, 2005: see below) and that attrition is purely due to a lack of offender motivation would seem to lack credibility. Instead it would seem that up to a quarter of all programme attrition is caused by the inappropriate implementation and

management of programmes. Such findings add weight to Nunes and Cortoni's (2006a, 2006b) calls for greater definitional clarity and observation of different types of programme dropouts both within research and practice.

Why is attrition a problem?

In a clinical sense, the main concern regarding attrition from programmes is obvious; those individuals who meet the criteria, and hence are identified as having a need for intervention, fail to benefit from the intervention. Considering this from the perspective of a medical model, if a patient failed to complete (or even start) their course of treatment, it is likely that this action would impact on the individual, in addition to having costs for society. The individual would not regain full health as easily as if they fully complied with their treatment, if at all. Such action could result in a need for further more extensive treatment which would have the consequence of reducing the resources available to others and increasing the costs to the health system. Within offender treatment, the effects are potentially similar. Assuming that the completion of treatment produces reductions in recidivism, the failure to complete (or start) the intervention has the potential to result in criminal behaviour and the associated costs of this that could have been prevented.

Such thinking is applied to offender interventions through the principles of risk and need posited by the risk-needs-responsivity (RNR) theory of Andrews and Bonta (2006) and operationalised within the programme accreditation criteria outlined above. The risk principle states that for programmes to be most effective the level of intervention, or dosage, should match the risk of reconviction of those subject to it. As such those offenders who pose a higher risk of future offending behaviour should receive a greater level of intervention. This theoretical principle receives research backing from studies which have compared low and high dosage programmes with low and high risk offenders. Generally, the results indicate that low risk offenders show larger gains in low intensity treatment and high risk individuals produce larger gains in high intensity treatment (Andrews & Dowden, 2006; Lowenkamp, Latessa, & Holsinger, 2006; Bonta, Wallace-Capretta, & Rooney, 2000).

The need principle suggests that interventions should be targeted at the needs of offenders that are related to their offending behaviour. As such they should be focussed

on dynamic criminogenic factors rather than on static or non-criminogenic factors. Andrews and Bonta (2006) draw on Dowden's (1998) review of treatment studies to create a list of criminogenic needs. These concur with Andrews and Bonta's 'Central Eight' (2006) risk factors which they argue are linked with increased recidivism: antisocial personality, antisocial attitudes and cognition, social support for crime, substance abuse, inappropriate parental modelling and discipline, problems in the school/work context, poor self-control, and a lack of pro-social activities.

As discussed above, the CSAP accreditation criteria specify that programmes should be clear in their targeting criteria so as to address the appropriate risk category of offenders (the risk principle) and should target dynamic criminogenic need (the need principle). It can therefore be assumed that, in line with Andrews and Bonta's principles, those offenders sentenced to attend a particular programme have been assessed as requiring the level and type of intervention that the full programme delivers. Considered in this context, attrition thus results in a violation of the risk and need principles and hence maximal treatment gains cannot be obtained. As Merrington and Stanley (2007) state:

> fulfilling the obligations imposed by a court order is clearly an important outcome for community supervision...But there is another reason for the importance of programme completion as an outcome measure. The 'treatment' benefits cannot be experienced if the offender does not attend and an offender who attends most of a programme should benefit more than an offender who attends only a little. (p. 442)

Evidence of the impact of attrition on outcomes (or indeed the lack of an impact on those who fail to complete) can be found within the literature. Within Hollis's (2007) large scale research, the Offender Group Reconviction Score (OGRS2; Taylor, 1999), calculated from criminal history variables to predict reconviction within two years, was compared against actual two year reconviction rates. Programme completers were seen to fare well: a reduction of 17% was observed between the predicted (64.3%) and actual rate (47.3%) of reconviction. Differences between predicted and actual rates were not observed, however, for non-starters and non-completers. Despite the absence of a

comparison group, this research demonstrates how non-attendance at an offending behaviour programme reduces the likelihood of treatment gains. Other research has reported similar findings; in considering the outcomes of treatment non-completers, Wormith and Olver (2002) found that at all levels of risk, but particularly for high risk offenders, reconviction rates were higher for non-completers than for completers. Such findings give credence to Underdown's (2001) statement: "Improving completion rates through higher offender compliance will be crucial to outcomes in a community context" (p. 118).

Other research has indicated, however, that the picture may not be as simple as this. McMurran and Theodosi (2007) conducted a meta-analysis of cognitive behavioural treatment studies comparing the recidivism rates of programme non-completers against untreated comparisons of comparable risk. Whilst completion of programmes was seen to be effective in reducing reoffending (d = 0.11), non-completion was found to be associated with *elevated* levels of reoffending (d = -0.16) with community samples of non-completers performing worse than prison samples (d = -0.23 and -0.15 respectively). Whilst such findings should be interpreted cautiously as the effect sizes reflect an averaging across studies (some of which reported reductions in reconviction for non-completers comparative to non-treatment comparisons) there is a suggestion within this work that dropout from programmes may not merely reduce the effect that completion would have produced but can instead be detrimental to programme outcomes. Similar results were reported by Hollin et al. (2008) in their study of community based general offending behaviour programmes. Programme non-completers were twice as likely and non-starters more than twice as likely as the comparison group to be reconvicted, even after controlling for any differences between the groups in terms of age, risk of reconviction, offence type, number of previous convictions, and follow up time.

Likewise, van Voorhis, et al. (2004) in their American study of parolees undertaking the R&R programme reported higher recidivism rates for non-completers compared to programme completers and a randomly allocated comparison group: within nine months, 60% of non-completers compared to 21% of completers and 40% of the comparison group had been rearrested or their parole revoked. Additionally, the time to recidivism was shorter for the programme non-completers compared to both

the programme completers and the comparison group: by the end of a six month follow up period 27% of the non-completers had been readmitted to prison whilst it was not until the end of the twelfth month that the comparison group reached the 25% re-admittance mark.

From this growing body of research, then, it could be concluded that attrition from programmes may in some way result in increased levels of reconviction. Such claims, however, should only be regarded as tentative. The majority of research within this area has compared the outcomes of programme dropouts against those of programme completers (e.g. Serin, Gobeil, & Preston, 2009) and/or a comparison group (e.g. van Voorhis et al., 2004). Gondolf (2001) warns against the use of programme completers as a suitable comparison group; if it is expected that a completed intervention has a positive impact then to utilise those on whom the impact is expected as a comparison to the programme dropouts would be weak research. An alternative comparison group is therefore needed. However, what represents a suitable comparison group has been a matter for negotiation and has formed part of a larger debate regarding appropriate methodologies for treatment evaluation.

Methodological issues in the study of attrition

The UK Home Office has argued that the evaluators of offending behaviour programmes should aspire towards the 'gold standard' of randomised control trials (RCTs). Within a Home Office published volume edited by Harper and Chitty (2005), a variety of authors proffer the message that: "Evaluations of correctional services interventions have often been based on sub-optimal research designs....to assess the impact of those interventions on re-offending, there is also a need to develop randomised control trials in the correctional services" (p. xx). Drawing on an adapted version of Sherman, Gottfredson, MacKenzie, Eck, Reuter and Bushway's (1997) Scientific Methods Scale, which classifies quantitative research according to its scientific rigor and places RCTs at the top of this scale, Harper and Chitty (2005) state that the RCT approach:

> ...minimises the chances that the treated and control groups differ in significant and important ways and that one group is biased from the

> outset to do better or worse. Failure to randomise means that studies are likely to include uncontrolled variables and/or selection effects, which mean interpretation of results is very difficult, if not impossible. (p. 7)

Other commentators in the field, however, have argued that RCTs within the criminal justice field are difficult to achieve in practice (Colledge, Collier, & Brand, 1999; Farrington, Gottfredson, Sherman, & Welsh, 2002; Gondolf, 2004; Hollin, 2008; Hollin & Palmer, 2009; Pawson & Tilley, 1997; Raynor, 2008) and could be considered unethical in that they withhold potentially beneficial treatment from a comparison group who, if it were not for the evaluation, would have received it (Hollin, 2006; 2008; Hollin & Palmer, 2009). Hollin (2008) has also argued that the results of good quality quasi-experimental evaluations (those which ensure tight controls between the experimental and comparison groups, either through matching on, or statistical control of, key variables) have not differed from those of RCTs. Indeed Heinsman and Shadish (1996) have concluded that if "Randomized and non-randomized experiments were equally well designed and executed, they would yield roughly the same effect size" (p. 162).Within a field where good quality randomised designs are problematic to implement and run, it would seem therefore that high quality quasi-experimental evaluation designs offer a viable alternative.

In relation to the study of programme evaluation, attrition and associated outcomes most quasi-experimental research designs have compared the outcomes of the comparison group against each of the naturally occurring sub-groups of the experimental group: completers and dropouts (Cann, Falshaw, Nugent, & Friendship, 2003; Hollin et al., 2004; Palmer et al., 2007; van Voorhis et al., 2004), or completers, non-completers and non-starters (Hollin et al., 2008; Hollis, 2007; McGuire et al., 2008; Stewart-Ong, Harsent, Roberts, Burnett, & Al-Attar, 2004; Roberts, 2004). However, in the absence of random assignment to treatment or comparison groups, so as not to undermine internal validity, methods for selection of an appropriate comparison group have to be considered. The first method that is used extensively within the field is the formation of a comparison group which is similar to the experimental group and then to statistically control for measured differences between the groups. Examples of such research are available in the literature: one such paper is that by Palmer et al. (2007)

which controlled for the influence of age, risk of reconviction, the number of previous convictions and the length of follow up period on the groups.

One of the limitations of such designs, however, is that statistical control can only be undertaken in relation to measured variables and the selection of these variables is the responsibility of the researcher. As such, it is perfectly possible that even after statistical control, the influence of a key unmeasured variable remains. Whilst tight control is therefore possible through careful selection of a comparison group and statistical control of any differences, quasi-experimental designs can never completely eradicate the possibility that an unmeasured variable is responsible for the outcomes observed.

Seager, Jellicoe and Dhaliwal (2004) argue that such comparisons are not valid for another reason: to compare the full comparison group to subsections of the experimental group (for example, the completers or the non-completers) fails to acknowledge potential differences between the subsections and hence the potential for subgroups within the comparison group. In reference to the sex offender treatment literature, the authors claim that previous evaluation research, in retaining the would-be non-completers within the comparison group sample but removing them from the treatment sample, invalidates the comparison group. When investigating treatment effects, the authors advocate the "purging" (p. 602) of the non-completion effect from the comparison group reconviction statistics to ensure a fair comparison:

> We think it is mandatory that all studies collect noncompleter data and then mathematically remove the calculated inflationary effect of noncompleters from the untreated comparison group. Failure to attend to this matter casts indefensible aspersions on any purported differences between the treated and untreated, whether reported in a single study or as a meta-analysis. (p. 609)

If, as Seager, Jellicoe and Dhaliwal advocate, to investigate a treatment effect the completers should be compared against that portion of the comparison group that would have completed the programme if they were given the chance, it therefore follows that research investigating the potential of a non-completion effect should do

likewise. The outcomes of programme non-completers should, therefore, be compared against those within the comparison group that would have commenced but failed to complete programmes. Furthermore, from such thinking it follows that the non-starters of programmes should always be considered within programme evaluation. If there are those within the comparison group that would commence but fail to complete a programme (if given the opportunity) then there are also those who would fail to commence at all. If, therefore, programme non-starters are not included within the research but their 'matches' are not removed from the comparison group, the evaluation of differences between the experimental and comparison groups is invalid (McConaghy, 1999; Nunes & Cortoni, 2006a).

One evaluation study which was able to compare completers, non-completers and non-starters with appropriate sub-sets of the comparison group reported on the community based Aggression Replacement Training (ART) programme for adult males (Hatcher, Palmer, McGuire, Hounsome, Bilby, & Hollin, 2008). The individuals within the experimental group were matched on a one to one basis utilising the variables of age, risk of reconviction and number of previous convictions. As such, it was possible to break the comparison group into subsets that were equivalent on these variables to the naturally occurring completers, non-completers and non-starters. The research reported a 7.8% increase in reconviction amongst programme non-completers compared to their matched comparisons and hence provided further tentative evidence of a non-completion effect.

Despite an ever growing swell of evidence, some commentators have argued that the differences in observed reconviction rates are not due to the impact (positive on the completers and negative on the dropouts) of the programme but instead reflect a process of self-selection. Debidin and Lovbakke (2005), for example, argue that unmeasured pre-programme differences between the completer and dropout groups result in both the dropout group failing to complete the programme *and* their elevated reconviction rates. As such, they claim that "the programme simply served to sort those who would do well anyway from those who would not, regardless of the treatment" (p. 48): those predisposed to a crime free future are those most likely to complete a programme. The corollary of this position is obvious: those who fail to start or complete are those most likely to recidivate. As such, these commentators argue that the

programme itself has no therapeutic value and serves purely to sift those who will desist from offending from those who will not.

Given that research to date has matched and/or statistically controlled for constructs such as age, risk of reconviction, and criminal history variables, if Debidin and Lovbakke's (2005) position is to stand, there must be some other unmeasured variable which influences both the propensity to dropout and to recidivate. Indeed, Debidin and Lovbakke claim that "completion rates are strongly linked to motivation" (p. 50) despite failing to specify what type of motivation they refer to and the distinct lack of evidence for such an assertion from the available literature. Seager, Jellicoe and Dhaliwal (2004), however, agree with such a position; in relation to their evaluation of a sex offender programme they stated: "that participation in the sex offender program did not reduce recidivism rates for those who complied with treatment but merely enabled motivated offenders to concretely demonstrate their commitment to not reoffend" (Seager, Jellicoe, & Dhaliwal, 2004, p. 609).

Hollin (2006) contests the view of Debidin and Lovbakke and argues that theirs is an unlikely explanation for the observed effects. Hollin (2006) proposes that the fact that a significant level of reconviction remains within the completion group invalidates the argument; it is not the case that all completers 'do well'. Hollin has also questioned the logic behind Debidin and Lovbakke's merging of the motivation and 'would do well anyway' arguments of completion; the motivation argument states that those who are motivated will complete a programme; whilst the 'would do well anyway' argument states that the programme is not an agent of change. Does this mean therefore that motivation is the agent of change? Hollin (2006; Hollin & Palmer, 2009) thinks that this is too simplistic an explanation: "If motivation is important in understanding desistence from further offending, it seems highly likely that it will take an interactive role with regard to other factors within the offender's life" (Hollin, 2006, p. 59).

As such, it becomes important to investigate the differences between the programme completers, non-completers and non-starters. Such knowledge will provide information to those working with offenders as to which are most likely to dropout but will also inform the programme evaluation literature. If key differences are discovered between the groups, subsequent quasi-experimental research could either statistically control for their influence or ensure that the experimental and comparison groups are

closely matched on this variable so as to rule out any potential influence it may have on treatment outcomes.

Correlates of attrition

In relation to the investigation of the correlates of attrition research has focussed on two main areas: those factors related to the offender themselves and those related to the programme or organisational procedures. These shall be considered in turn.

Individual Factors

International research across a variety of offending behaviour programmes has established that programme dropouts tend to have higher risk of reconviction scores (Craissati & Beech, 2001; Browne, Foreman, & Middleton, 1998; Nunes & Cortonni, 2006a; Turner, 2006; Wormith & Olver, 2002), more previous convictions (Babcock & Steiner, 1999; Zanis, Mulvaney, Coviello, Alterman, Savitz & Thomson, 2003), are more likely to have a previous violent conviction (van Voorhis et al., 2004) and tend to be younger (Hazeltine et al, 2002; Mosher & Phillips, 2001; Nunes & Cortoni, 2006a; Robinson, 1995; Turner, 2006; van Voorhis et al,. 2004; Zanis et al., 2003) than programme completers. In addition, programme dropouts tend to be less well educated (Babcock & Steiner, 1999; van Voorhis et al. 2004; Wormith & Olver, 2002), and less stable within their lives (Butzin, Saum, & Scarpitti, 2002; Craissati & Beech, 2001; Roberts, 2004; Wormith & Olver, 2002).

Within research from England and Wales, a similar pattern emerges. Palmer et al. (2007) reported that community based general offending behaviour programme dropouts (non-completers and non-starters within one combined sample) were significantly younger, higher risk, and had more previous convictions than the programme completers within their sample. These findings have been supported by other research projects within England and Wales (Chopourian, 2003; Gill, 2004; Hollin et al. 2008; Hollis, 2007; Roberts, 2004).

Programme dropouts have also been found to have more risk factors or criminogenic needs than programme completers. For example, Chopourian (2003)

undertook an analysis of completers and dropouts of the Think First programme within the West Midlands Probation Area and found that more dropouts than completers were found to have problems within the areas of education and employment, finances, lifestyle and associates, thinking and behaviour, and attitudes to offending as assessed by the Offender Assessment System (OASys[2]). Additionally, dropouts had more problems in their lives and were more cognitively impulsive than the completers. However, the sample size utilised within this research was small, the quality of the research was basic, and the dropouts were not separated into non-completers and non-starters.

Other research, however, has reinforced the findings of Chopourian (2003), especially where dropout was related to withdrawal or expulsion from the programme, as opposed to administrative reasons or personal circumstance. For example, Nunes and Cortoni (2006a) in their study of nearly 8000 offenders sentenced to a correctional programme found such dropouts to have greater criminogenic need, and lower levels of motivation for intervention than the completers and the administrative non-completers whilst comparisons between all other groups did not reveal any differences in relation to risk, need, or motivation. The aim of Nunes and Cortoni's project was admirable: to determine whether different types of programme non-completers have different correlates and hence reasons for dropout. However, the research suffers from being rather too ambitious: the sample of nearly 8000 offenders across Canada included those held within secure establishments and those within the community, on a range of different programmes with a range of intensities. These factors were not, however, disaggregated by the researchers and hence it could be argued that this research suffers from the same criticisms that the author themselves level at others: it attempts to find solutions to a problem with a heterogeneous group. The research could have been improved if the different types of attrition within the different programmes and settings had been investigated. Indeed in an associated research project, Nunes and Cortoni (2006b) reported differing rates of attrition across the different programmes types and intensities indicating that either the types of offenders varied across programmes and/or that attrition may not derive solely within the offender.

[2] OASys is the joint prison and probation risk and need assessment tool utilised within England and Wales.

Using a mixed methodology, Westmarland et al. (n.d.) undertook to examine correlates of attrition within the Northumbria Probation Area. The research analysed PSR reports of those recommended to attend the Think First programme, the criminal histories of those sentenced to the programme, the post-programme evaluation forms of programme completers, and conducted interviews with case managers, programme completers, and programme non-completers[3]. From the collected data, the authors were able to determine that programme completers were more likely than non-completers and non-starters to have problems with stress management and alcohol and/or drugs and less likely to have issues with boredom/need for excitement or with self-esteem/self-image. Non-completers had more problems than completers and nonstarters in relation to reasoning/thinking skills, inter-personal/social skills, mental health, emotional well-being, and learning disabilities and/or literacy problems. The authors take this as evidence that this group "were prepared to attempt the programme but found it difficult to complete" (p. 16). Non-starters were more likely than the non-completers and completers to have criminal peers and less likely to take responsibility for their offending behaviour and recognise the harm caused to their victims. It would appear, hence, that the non-starter group were more criminogenic in their nature and less willing to take responsibility for their actions than both the completer and non-completer groups.

Another individual factor that has been implicated in its influence on attrition is the presence of literacy problems (Briggs, Gray, & Stephens, 2004). Davies, Lewis, Byatt, Purvis, and Cole (2004) undertook a review of the literature demands of the three UK general offending behaviour groupwork programmes and concluded that there was a mismatch between the literacy levels required by the programmes and those held by the offenders referred to them. Roberts (2004) found that this mismatch may impact on attrition: completers were found to have better verbal communication and literacy skills that the non-completers. Wormith and Olver (2002) also found that programme completion was linked to educational level and discussed whether it was necessary therefore to alter the programme for various needs or provide additional pre-

[3] It should be noted, however, that access to non-completers was limited and, hence, a proportion of 'non-completers' were taken from a pool of programme re-starters and perhaps, therefore, actually represent a subset of non-completers or indeed completers should they go on to complete the programme upon which they have restarted.

programme support and preparation. Whichever is the best solution, Wormith and Olver's findings suggestions infer that it is the responsibility of the correctional service to pay greater attention to individual issues to reduce attrition rates.

The study of the contribution of motivation, or indeed a lack of it, to programme attrition has revealed contrary results. As discussed above, Debidin and Lovbakke (2005) claim that the concept of motivation is strongly associated with programme completion and attrition. In line with this perspective, Scott (2004) found that motivation to change predicted dropout from a domestic violence treatment programme. Motivation for change has also been found to predict dropout and expulsion among correctional samples (Beyko & Wong, 2005; Krawczyk, Witte, Gordon, Wong, & Wormith, 2002, cited in Nunes & Cortoni, 2006b; Mckenzie, Witte, Beyko, Wong, Olver, & Wormith, 2002, cited in Nunes & Cortoni, 2006b; Nunes & Cortoni, 2006a; Wormith & Olver, 2002) and within the substance use treatment literature (De Leon & Jainchill, 1986; De Leon, Melnick, Thomas, Kressel, & Wexler, 2000; Ryan, Plant, & O'Malley, 1996; Simpson & Joe, 1993; Simpson, Joe, & Rowan-Szal, 1997).

In relation to community based general offending behaviour programmes, Al-Attar (2003) claims that offender motivation is required, albeit alongside an understanding of the programme's aims, for a programme to be effective and treatment gains to be maintained (however, in stating such claims, Al-Attar provides no evidence from research for these statements). Wormith and Olver (2002) also reported that motivation was one of a number of factors upon which treatment completers and non-completers were found to differ. McMurran and McCulloch (2007) also reported higher self–reported motivation scores amongst prison based offender behaviour programmes completers than non-completers (77.5% and 66.5% respectively). It is possible, given the retrospective nature of this research, however, that these ratings may have been affected by subsequent events and that the non-completers rated their motivation as lower in part due to their non-completion status.

On first inspection, this body of evidence may seem somewhat persuasive. However, as Hollin comments, "the main problem with the unqualified use of the term "motivation" is the lack of precision in its meaning" (2006, p. 59). Drieschner, Lammers, and van der Staak (2004) agree and claim that within treatment motivation research, the behavioural outcome relating to the motivation (i.e. treatment completion, crime

desistance) is, at best, not clarified and, at worst, results in a circular argument whereby a lack of motivation to engage in treatment is demonstrated within research by programme dropout but is also used as an explanation for this dropout. In reviewing the papers above, these concerns would appear justified: the term 'motivation' is rarely defined and the behavioural outcome to which it refers is often not clear. Even in cases where the behavioural outcome is defined, there is often confusion as to what this means. For example, research by Brocato and Wagner (2008) on attrition from a substance misuse treatment programme found that problem recognition predicted the number of days that the client remained in treatment. The authors take this finding to suggest that those who were more 'motivated to change their problem behaviour' or 'treatment ready' at the time of entry into the programme remained in treatment longer. It is not clear, however, how greater problem recognition is related to motivation to change: just because a problem is recognised does not mean that the individual wishes to change it. Additionally, where research has attempted to measure 'motivation' directly the researchers often utilise staff ratings of the offenders' participation in the treatment. Given the claim by Hanson and Bussiere (1998) that "Motivation to change is difficult to assess…because there are clear benefits to "appearing" willing to change" (p. 349) caution should be urged when assessing the implications of these results.

It is not surprising that research assessing the correlates of programme completion and attrition has attempted to use this information to devise tools to identify those at risk of dropout. Nunes and Cortoni (2006b) compared programme completers with those who were withdrawn or expelled from treatment on a range of risk, need, and motivation factors and used those that differentiated the groups (a measure of risk, age, marital/family need, prosocial attitudes, and motivation for intervention) to form the Dropout Risk Screen. The tool "predicted dropout/expulsion with a moderate and a statistically significant degree of accuracy" (p. ii). The authors warn, however, that such a tool should not be used to determine how treatable an offender is but should instead trigger further assessment and pre-treatment work with those identified at a higher risk of dropout.

Attempts have also been made within Australia to develop a tool which measures "treatment readiness" (Casey, Day, Howells & Ward, 2007, p. 1427). Drawing

on the theoretical work of Ward, Day, Howells and Brigden (2004), which proposes the influence of internal and external factors on treatment engagement, Casey and colleagues have developed a tool that measures the internal characteristics of treatment readiness so as:

> to inform professional assessments of readiness in a way that is both cost- and time-effective and might be particularly valuable in the routine assessment of suitability for programs such as cognitive skills, which may be offered to large numbers of offenders. (p. 1428)

The Corrections Victoria Treatment Readiness Questionnaire (CVTRQ) is a self-report tool which measures four components: attitudes and motivation towards programmes, emotional reactions to the individual's offending behaviour, offending beliefs, and efficacy relating to the individual's perceived ability to engage in their treatment. Whilst the CVTRQ has been found to be associated with other measures of treatment engagement (Casey et al., 2007), it is yet to be tested as a predictor of programme attrition.

Beyko and Wong (2005) concur with Nunes and Cortoni (2006b) in arguing that attrition profiles should not be used to exclude individuals from treatment. Instead they argue that such information should serve as warnings to the service providers that the treatment provision does not match their clientele: "Treatment attrition can be seen as our clients voting on our services with their feet. Unless we listen closely to how our voters vote, we may end up totally alienating our constituents" (p. 388). As such, the authors acknowledge that individual factors can and do contribute to attrition but see the resolution of this as being the responsibility of the service provider rather than the offender. Such thinking concurs with Andrews and Bonta's (2006) third principle of effective interventions. In addition to the principles of risk and need (discussed earlier), Andrews and Bonta's proposed the 'responsivity' principle which states that the programme design and delivery style should match to target offenders' learning styles so as to increase their engagement in the programme. Andrews and Bonta argue that general responsivity, which relates to the delivery model, is straightforward in that programme development should be guided by research findings. Specific responsivity

which relates to how the delivery of the intervention is matched to the learning styles of individual offenders, however, can be more difficult to ensure and remains under-researched (Andrews, Bonta, & Wormith, 2006). However, Andrews and Bonta do argue that, in order to conform to the principles of risk and need, correctional staff should consider those factors within the individual that may be associated with low engagement or programme dropout and should aim to address these so as to increase engagement.

Within Sussex Probation Service's (2003) report on attrition, however, was an implicit message which advocated a 'cherry picking' of those offenders most likely to complete programmes ahead of those who might require more support throughout the treatment process. Indeed one suggestion to increase completion rates was to "Target those likely to complete and fast track them on to a course (this is being experimented with)" (p. 17). Such thinking, however inevitably leads to resources being diverted to those offenders who would not benefit from intervention and away from those who need it most. Palmer, McGuire, Hatcher, Hounsome, Bilby and Hollin (2008; 2009) reported that those who are most likely to complete a programme are offenders who are at low risk of reconviction; however, it is this same group who see the smallest benefits from completion coupled with a large negative effect of non-completion. As such, targeting those most likely to complete may result in more favourable attrition rates but it is unlikely that these will translate into reduced reconviction rates. Such solutions by Probation Areas are not surprising, however, given that they are working towards ambitious funding-related completion targets; in this context, Probation Areas are faced with the question as to whether the extra investment of supporting a 'difficult' offender through a programme is worth the gain of one completion towards a target of a few hundred?

Whilst further research with attrition or low engagement prediction tools would add to the knowledge base in relation to attrition, the CVTRQ, and other such tools, fail to consider organisational or programme factors that may contribute to attrition. As Hollin and Palmer (2009) state

> The level of programme completion (and hence the completion effect) must be, in part at least, a consequence of programme design and style of

> delivery. Programmes high in responsivity will engage offenders, thereby increasing completion rates and, arguably, reducing reoffending. (p. 159)

The following section will assess the research literature that has addressed the contribution of such variables to the concept of attrition.

Organisational, Process, or Programme Factors

Within the community services of England and Wales, the contribution of process factors to attrition has been acknowledged by the National Probation Directorate (NPD). The NPD have indicated concern relating to three issues: first, attrition rates generally but especially those of the high risk offenders; second, the under-representation of medium risk offenders referred to offending behaviour programmes; and third, the inconsistency between the risk profile of those referred and the targeting criteria outlined within offending behaviour programme manuals (NPD; 2002). Amid such concern Kemshall and Canton (2002) were commissioned to "identify good practice for the reduction of attrition pre-programme, with particular attention to how to reduce drop-out between Court Order and programme commencement" (p. 4). All probation areas were asked to identify the main reasons for pre-programme drop out from offending behaviour programmes within their area. The most cited reasons were: a lack of offender motivation, poor pre-programme preparation, waiting times, access and transport issues, case management, inappropriate targeting, breach for further offences, and breach for non-attendance.

Two broad conclusions can be drawn from this research. First, the majority of the factors in the list of most cited reasons for pre-programme attrition relate to implementation or organisational issues; only 'breach for further offences', 'breach for non-attendance' and 'a lack of offender motivation' infer offender related factors as the main contributor to attrition (and even so it could be argued that non-attendance could be an outcome of the other listed factors and that motivation should be enhanced by good case management). Second, the recording of reasons for dropout from programmes by Probation Areas is poor. Examination of the reasons provided to Kemshall and Canton, reveal that many are vague and provide little information as to

why someone did not commence a programme. To know that someone failed to commence for the reason of 'breach for non-attendance' tells us only that the individual did not attend a probation session of some description; it is not clear whether this is a programme session, a case management session, or another type of session. It also does not tell us why they failed to attend this session. Likewise, if 'inappropriate targeting' is given as a reason for attrition, does this mean that the targeting was inappropriate because it focussed on individuals of the wrong risk category, or with inappropriate needs, or because the procedure had referred appropriate but unsuitable offenders, for example, those that lacked motivation to attend? Much more research is required to unpick these issues.

Similar conclusions can be extracted from Briggs and Turner's (2003) and Stephens and Turner's (2004) classifications of programme non-starters into Stephen's action oriented typology of reasons for failure to start (Stephens, 2003). After three months, only 25% (Briggs & Turner, 2003) and 30% (Stephens & Turner, 2004) of the samples had commenced a programme. Of the remainder, in both studies almost ten percent were classified as type one offenders; they were not able to commence a programme due to being in custody, hospital, moving area, or having the condition removed from their order. In the Briggs and Turner (2003) paper, the proportion of type two offenders, those available to commence but unsuitable for a programme, was low at just one percent of the sample. The report by Stephens and Turner (2004) using data from 2002[4], however, had previously placed this figure at six percent. The Probation Area claims that in the time period between the two studies, that targeting and assessment practices had been improved and that this had successfully impacted on type two attrition.

Despite such a reduction in type two attrition, 65% of the sample were classified as type three offenders (up from 56% in the Stephens and Turner research); 16% did not have a group place available to them as groups were full, or had been inappropriately scheduled to fit with curfews, employment or childcare responsibilities; 24% had not commenced a programme because they were already subject to breach proceedings,

[4] Although the publication dates indicate otherwise, the Briggs and Turner research took place after the Stephen's and Turner research and hence perhaps indicates that changes implemented by the Probation Area following the Stephens and Turner research has improved the rate of Type Two offenders.

either due to missing case management or pre-programme appointments; and a further 25% were classified as available, suitable, with a group place available and not in breach but had still not commenced. Echoing Kemshall and Canton's (2002) research conclusions, reasons for this latter group failing to commence their programmes were not available due to a lack of data recording. What is clear, however, is that a significant proportion of pre-programme dropout was influenced by process rather than individual factors and that, as was seen in the case of targeting and assessment practices, the alteration of organisational practices can have an impact on attrition rates.

There is also clear evidence within the academic literature of the contribution of inappropriate targeting and allocation to attrition. Palmer et al. (2008; 2009) have demonstrated the variation in targeting practices between probation areas. The rate of 'appropriate' allocation, defined as when the offenders' risk of reconviction score matched the specified targeting criteria within programme manuals, varied between probation areas from 37.1% to 82.1%, with almost half of all offenders falling outside the recommended risk of reconviction banding[5]. Further verification of such misallocation comes from Turner (2006) who reported that a quarter of ETS programme participants within West Yorkshire had risk scores below that of the targeting criteria. Westmarland et al. (n.d.) found that case managers felt that such inappropriate allocation to programmes occurs due to the lack of programme knowledge amongst pre-sentence report writers and the pressure to recommend offenders to meet performance targets. This paper also found that case managers believed that such practice, rather than contributing to meeting such targets, actually increased attrition rates. The authors quote one of the case managers in relation to the latter point: "there is no point in trying to get more and more people recommended for the course when you know that a lot of them are not up to it. You're just setting them up to fail – there is a moral issue here" (p. 23).

Gill (2004) also found that half of the non-completers within a local sample of offenders referred to the Think First programme had risk of reconviction scores above the upper limit whilst only one of the programme completers fell above this range. Within Palmer's work, which was drawn from a national data sample, eight percent of

[5] The manuals of the general offending behaviour programmes at that time specified an OGRS2 score of between 31 and 74 (National Probation Directorate, 2001).

the sample had risk of reconviction scores which fell below the targeting criteria, 52% were considered to fall within the appropriate banding, whilst just short of 40% were considered to be too high risk to be targeted for the programmes under consideration. With completion rates varying with risk such that high risk offenders are the least likely group to complete programmes (23.7% compared to 35.5% from the appropriate risk category and 50% of the too low risk category (Palmer et al., 2008)), the finding that nearly 40% of those allocated fell within the 'too high' category lends weight to the argument that such organisational factors do contribute to attrition rates.

Westmarland et al. (n.d.) in their local report took the approach of asking case managers their views of the factors associated with the non-starting and non-completion of a programme. The responses received fell into two pools: those external and those internal to the programme. The factors mentioned by case managers as being external to the programme were: the offenders' anxiety about going on a programme, a lack of offender motivation, literacy problems, mental health/emotional problems, substance use and the associated chaotic lifestyles, re-offending, a lack of accommodation or childcare provision, reoffending, and the gaining of employment. Whilst the factors provided by case managers here related to offender factors, they reported that appropriate identification of these issues at pre-sentence report stage would allow the case managers to support the offenders through their programme to completion. Such thinking would concur with the responsivity principle of effective offender interventions (Andrews & Bonta, 2006): organisational factors can influence offender compliance and hence outcomes.

Evidence from Wormith and Olver's (2002) research suggests that a lack of responsivity does indeed impact on attrition rates. The authors found that only 20% of very high risk aboriginal offenders completed their evaluated programme compared to 67% of equally high risk non-aboriginal offenders. Wormith and Olver claim that such findings highlight concerns about the ability of a programme (or the staff delivering it) to accommodate responsivity issues inherent within the aboriginal culture. There is therefore a call for correctional organisations to match the provision of services to offender needs to ensure maximal engagement. Returning to the findings of Westmarland et al., appropriate targeting coupled with methods to alleviate offender anxiety, flexibility of programme scheduling to account for employment,

accommodation and childcare issues, and pre-programme work with offenders to increase their motivation to cease offending and attend their intervention would, it would seem, increase programme engagement. Such an approach, however, necessitates holistic programme training with pre-sentence report writers and case managers requiring instruction on how to identify obstacles to engagement and how to work with the offender to increase the chances of their cooperation with their sentence.

Turner (2006) has suggested that improvements in staff training, both programme facilitator and case management, should be made in an attempt to increase within-programme retention. Within his research, Turner interviewed probation staff who commented that the accredited programme training they had received had prepared them well for the technical aspects of the programme. However, roughly half of the staff sample did not feel adequately skilled to deliver the programmes for which they were trained. The staff called for more training in relation to group work skills and group dynamics to aid their ability to be responsive to group members' learning styles. In relation to case management, Turner devised a case management programme integration checklist and discovered that the case management received by programme completers and non-completers differed on aspects of this measure. Programme completers were more likely to receive positive in-programme support from their case managers and more organised case administration. Whilst it is not possible from this research to infer causality (it is just as likely that a lack of offender programme engagement influenced the case management practice as the alternative), this research provides an indication that variations in case management provision, which may relate to training issues, are associated with treatment attrition.

Sussex Probation Area (2003) also highlighted training issues relating to case management that may impact on attrition. This research found that case managers were unsure when they were required to deliver the pre-programme sessions that comprise the Think First programme, that they were frustrated by long periods between order and programme start, and felt ill-equipped both in relation to training and resources. Such findings seem to concur with the opinion of Kemshall and Canton (2002) that "The level, content and quality of engagement with offenders pre-programme is crucial to attrition" (p. 5). Similar findings were reported within a process evaluation commissioned by the Home Office prior to the widespread implementation of

programmes within the Probation Service (Hollin et al. 2002a, 2002b). It would seem therefore that this issue has still not been adequately addressed by senior management.

Such findings appear to indicate that the integrity of the programme may have been undermined by the fast pace of implementation and that this in turn has contributed to low engagement and offender dropout. The term 'programme integrity', originally coined by Quay (1987), refers to the delivery of the programme as intended by theory and design. Hollin (1995, p. 197) claims that there are three threats to programme integrity: "programme drift", whereby the aims of the programme are altered gradually over time; "programme reversal", whereby the aims of the programme are undermined by the staff through the use of procedures or the modelling of behaviours that work against the programme philosophy; and "programme non-compliance", whereby the programme deliverer opts to change or omit particular sessions of the programme. It is likely that insufficient or inappropriate training throughout the service could contribute to a lack of integrity: if the staff do not understand or appreciate the theoretical basis of the programme, it perhaps becomes more likely that a threat to its integrity could occur.

Turner (2006) reported that the integrity of the programmes within his locality had, at times, been judged to be 'poor' or 'very poor': programme deliverers reported insufficient time to prepare, debrief and review their sessions, staff shortages, a lack of groupwork skills, and inexperienced or poor treatment managers (those responsible for managing programme integrity issues). Over a third of staff also commented that they were either "not at all" or only "a little" supported by their treatment manager or programme manager: they called for more treatment management and easier access to these individuals. Given the work by Lowenkamp, Latessa, and Lemke (2006), Lowenkamp, Latessa, and Smith (2006) and Feldman and Wodarski (1983, cited in Wormith & Olver, 2002) which has demonstrated poorer outcomes for programmes with low integrity compared to those which adhere to the principles of effective practice, it is possible that the organisational factors reported by Turner also have their bearing on programme attrition.

Other factors that the literature indicates may impact on attrition include the length of time between receipt of the probation order and programme commencement (Chu, 2003; Turner, 2006), previous breach of a probation order (Chu, 2003), the lack of

evening provision (Sussex Probation Area, 2003), the "chopping and changing of tutors" (Sussex Probation Area, 2003, p. 12), the over-rigid lateness and missed session rules (Westmarland et al., n.d.), the demotivating impact of pre-programme psychometric booklet completion (Westmarland et al., n.d.), and the contents of the early sessions of the programme (Sussex Probation Area, 2003; Ong, Al Attar, Roberts & Harsent, 2003). Further research is required in relation to such factors, however, as the evidence for such connections is, thus far, merely indicative.

Offender perceptions of attrition

One of the few studies to date which has asked offenders directly about their experiences of dropout from offending behaviour programmes was conducted with offenders who had failed to complete the ETS programme within institutional settings (McMurran & McCulloch, 2007). Of those interviewed, all programme dropouts stated that they had negative feelings about failing to complete the programme. Half of the sample stated that the intervention met their needs but had still failed to complete, whilst the other half stated that the programme was not right for them. The reasons provided for this perceived mismatch between themselves and the programme was diverse: some offenders commented on factors relating to the programme itself stating that it was patronising and too simplistic, whilst others reported that they had not received adequate support throughout the programme. The sample, however, contained a mix of those who had been removed from the programme, those who chose to withdraw, those who had to withdrawn due to ill health, and those who had been released from prison during the run of the programme and hence were not able to complete. The results were not separated according to the type of non-completion and hence given the findings of Pelissier, Camp, and Motivans (2003) concerning the different correlates of different types of attrition, the conclusions that can be drawn are limited. In addition, it is likely that a sample undertaking treatment within the community is likely to yield some very different reasons for non-completion. It would therefore be interesting to investigate the factors associated with attrition amongst the different types of dropouts amongst a community based sample.

Wood (2005) carried out a small sample qualitative study on community based offenders who had failed to provide an acceptable reason for not commencing a

programme (non-starters). These offenders were not able to specify how the programme may have been beneficial to them despite having attended pre-programme preparatory sessions (however, it was not assessed whether programme attendees were able to do this and hence this may not be particular to programme non-starters). The vast majority of offenders reported that they had agreed to the programme at the sentencing stage in order to avoid custody. However, once they learned of the commitment involved in the intervention they felt that the sentence was too severe. Wood hypothesises that such opinions have arisen due to the belief amongst offenders that probation is a soft option. She argues that such offenders are yet to appreciate the philosophical shift within probation to the punitive correctional perspective and hence do not appreciate the consequences of non-compliance. This proposed explanation is countered, however, by the finding that offenders decided not to attend the programme despite understanding the consequences of their decision.

Whilst Wood's research makes an interesting contribution to the knowledge base, it is not possible to state with any certainty that the reported factors are unique to programme non-starters. In the absence of programme attendees within the sample, it is not possible to determine whether links between the above factors and attrition are valid. Such research would be strengthened by seeking the views of programme non-completers and completers, in addition to those of programme non-starters, to determine if these issues arise with equal emphasis amongst these groups also.

Combining individual and organisational/process factors

From the above body of literature assessing the correlates of attrition, it would seem that there are indeed individual and organisational factors that impact on programme attrition. In recognition of this, authors such as Ward et al. (2004), Serin (1998) and Serin and Kennedy (1997) have been instrumental in the development of the concept of 'treatment readiness'. The early work of Serin and colleagues has provided eleven indices, internal to the offender, which contribute to offender readiness for treatment: problem recognition, goal setting, motivation, self-appraisal, expectations, behavioural consistency, views about treatment, self-efficacy, dissonance, external supports and affect (Serin, 1998). Serin proposes that the offender's treatment readiness, rated as low, medium, or high based on the presence or absence of these

indices, can be combined with ratings of treatment performance (which are seen to be influenced by factors relating to the setting of the intervention) and assessments of risk to determine how responsive the offender will have been to the treatment undertaken. As such, the treatment readiness information is seen as useful to determine the likely *impact* of completed treatment on the offender and hence the potential for their movement from secure accommodation to the community, rather than as an indication of whether the offender will *engage* with a future intervention.

The concept of 'treatment readiness' as defined by Ward and colleagues is:

> the presence of characteristics (states or dispositions) within either the client or the therapeutic situation, which are likely to promote engagement in therapy and that, thereby, are likely to enhance therapeutic change.... To be ready for treatment means that the person is motivated (i.e. wants to, has the will to), is able to respond appropriately (i.e. perceives he or she can), finds it relevant and meaningful (i.e. can engage), and has the capacities (i.e. is able) to successfully enter treatment (p. 647).

As such, the model, entitled the Multifactor Offender Readiness Model or MORM, focuses on dimensions internal to the offender – cognitive (attitudes, beliefs, expectations, self-efficacy), affective (level of emotional control, distress, guilt, shame), behavioural (evaluate behaviour as a problem, seek help, have competency to participate in therapeutic process), volitional (intention, motivation to change, personal goals compatible with rehabilitative goals), and personal/social identity factors – in addition to those external to the offender – circumstances (perceived level of coercion and the offenders response to this), location (community or prison), opportunities (availability of treatment, stage of sentence), resources (to deliver goo quality programmes), support (family, holistic organisational support, rewards) and programme/timing factors (appropriate treatment at an appropriate time). The MORM proposes that as these factors combine, the likelihood that the offender is ready to engage in treatment increases. As such, the authors propose that the MORM can "serve as a heuristic model for clinical decision making" (p. 665-666) when working with offenders who are resistant to treatment. This may involve modifying the client, the

treatment programme, or the setting in order to increase readiness and hence offender engagement. As such, the authors argue that "it is likely that offenders with low readiness across multiple areas will be those that either do not complete treatment or for whom treatment does not impact on their offending" (Ward et al., 2004, p. 668).

The proponents of the MORM thus concur with the argument of Beyko and Wong regarding how services should use this information. Rather than use indications of treatment readiness as selection tools for who should and should not receive treatment, Day and colleagues conceptualise treatment readiness as adaptable and the responsibility of the service provider:

> the primary responsibility for modifying low levels of readiness lies with the practitioner and the service in which s/he works. In other words, it becomes the treatment provider's duty to find ways to deliver interventions that meet the individual's needs, and discourages the tendency to view low levels of motivation or compliance as pathological (Day, Howells, Casey, Ward, & Birgden, 2007, p. 22).

A further quote by taken from Day et al. (2007) highlights how the concept of treatment readiness can link with the study of treatment attrition: "a consequence of being 'ready for treatment' is the ability to not only attend rehabilitation programmes, but also to engage with the programme content, facilitators, and other group members" (p. 23). Thus, a person who is treatment ready will first attend and then engage with treatment. As such, in an attempt to learn more about the factors that impact treatment readiness is becomes necessary, first to study those variables, both internal and external to the offender, which impact on treatment attendance followed by those that influence engagement with treatment attended. Research by Langevin (2006) has reported that differences do indeed exist between those who claim to want treatment but do not attend, those who attend treatment but do not complete, and treatment completers. As such, it would seem likely that there are different readiness factors associated with programme attendance and programme engagement. This thesis therefore aims to evaluate the first of these issues. It will assess those intrinsic and extrinsic factors associated with attendance at, and completion of, rehabilitation

programmes. Further research should build upon this work to determine those factors that influence offenders who attend the intervention to subsequently engage with the programme contents. Of course, one indication of engagement is whether the offenders complete the programme to which they are sentenced and hence in evaluating the non-starters, non-completers and completers of interventions, this thesis also touches on the issue of engagement. However, programme completion is not the only indicator of engagement: it is perfectly possible for an offender to complete the programme to which s/he is sentenced without engaging directly with, or benefitting from, the programme. Thus, the combination of work assessing the correlates of, and reasons for, programme attrition with that investigating the factors associated with treatment engagement will inform the treatment readiness, programme effectiveness, and treatment development literature, and hence feed into clinical practice with offenders.

Aim of this research

The aim of this thesis, therefore, is to advance knowledge in relation to the factors associated with programme attrition within a sample of offenders sentenced to a general offending behaviour programme within the community services within England and Wales. In recognition of the need to differentiate between different types of attrition, this research will assess for differences between programme completers, non-completers, and non-starters of such programmes. It is hoped that this information will not only provide information which will influence correctional services' practices in relation to programme and offender management but will also inform the treatment readiness and effectiveness debate: the more that is known about programme attrition – who drops out, who completes, the correlates and reasons for dropout and completion – the more we can learn about how such factors influence programme attendance and outcomes and how we can work to improve completion and recidivism rates.

PART ONE

Investigation of attrition and reconviction outcomes utilising national probation data

An Introduction to Part One

Part one of this thesis aims to investigate attrition from two community based offending behaviour programmes (Enhanced Thinking Skills and Think First) utilising a national sample of offenders.

The focus within this section of the thesis is twofold. First, this section will conduct an evaluation of the outcomes of the programme completers, non-completers and non-starters relative to a comparison group which has been matched on a one to one basis. Initially the completers, non-completers, and non-starters will be combined into an experimental group to determine whether there are any differences in reconviction between this and the comparison group. Then the naturally occurring groups of completers, non-completers, and non-starters will be examined relative to their matched comparisons.

Second, this section will report on an examination of differences between the three groups of programme completers, non-completers and non-starters. The groups will be compared using demographic (age, literacy, school leaving age, employment, accommodation), criminogenic (risk of reconviction, number of previous convictions), intervention related (suitability, motivation) and offender need related (total, criminogenic and non-criminogenic) variables. Where differences are determined, the impact of these on the relationship between completion, dropout, and subsequent reconviction will be investigated.

Chapter Two

Attrition: An analysis of reconviction outcomes

Introduction

The What Works question, on the face of it, seems straightforward: does offender programming reduce reconviction outcomes? However, practical and methodological issues encountered by researchers attempting to answer this question ensure that it is not so straightforward after all. One of the many dilemmas facing evaluation researchers is how to deal with the inevitable attrition from offending behaviour programmes within the research design and analysis. Not only can attrition be significant in magnitude (Hazeltine et al., 2003; Hollin et al., 2002a; 2004; 2008; Hollis, 2007; Kemshall & Canton, 2002; Westermarland et al., nd), it also presents methodological problems: should research consider the outcomes of the programme's dropouts in addition to those of the programme's completers? If so, during the analysis should programme dropouts be included with the programme completers within an 'experimental' group, or, should dropouts be investigated separately from the programme completers?

Taking the first of these questions, it can be argued that programme dropouts are as much a product of the intervention as programme completers and hence should be considered within an evaluation. As Al-Attar (2002) states:

> if we only examine the programme effectiveness for completers, we are basing our conclusions on a biased, selective sample… As this sample is not representative of the total sample allocated to the programme, the conclusions about effectiveness are then limited in their generalisibility. (p. 131)

Some of the previous intervention evaluation research has used the programme dropout group as the comparison against which the programme completers are evaluated (Gondolf, 2001). As Gondolf argues, such methodology is ineffectual: programme completers may have favourable outcomes in comparison to the programme dropouts "simply because they are less prone" (p. 80) to unfavourable outcomes. Any observed differences between the groups, hence, could not be

considered evidence of a treatment completion effect: it could be that programme completers are those that "would do well anyway" (Debidin & Lovbakke, 2005; p. 48) or indeed any differences could be due to the impact of dropout rather than the impact of programme completion. As such, this design is not helpful to those wishing to establish whether offender programming can reduce reconviction or recidivism outcomes.

An 'intention to treat' or 'treatment allocated' design, is one that includes the programme completers and dropouts together within the 'experimental' group and compares the outcomes of this group with those of an appropriate group of comparisons. Such a design is usually reserved for evaluations utilising randomisation in their allocation to the experimental and comparison groups. Whilst providing a reasonable design for an evaluation of *rehabilitative policy*, the problem with this design for use within a *treatment evaluation*, however, is that the programme dropouts within the experimental group have not received the full dosage (if any) of the treatment. As such, and in line with RNR theory (Andrews & Bonta, 2006), it is not expected that this group of offenders would have benefitted from treatment, having been exposed to only some or, indeed, none of the content. As such, any potential effect of programme completion would be diluted by the outcomes of the programme dropouts (Hollin, et al., 2008; McGuire et al., 2008). In cases where dropout is high (such as within community settings in England and Wales (Hazeltine et al., 2003; Hollin et al., 2002a, 2004, 2008; Hollis, 2007; Kemshall & Canton, 2002; Westermarland et al., nd)), any completion effect would be swamped by the outcomes of the programme dropouts.

Such findings have been observed amongst the work of a number of offending behaviour programme evaluators. van Voorhis et al. (2004) in their evaluation of the Georgia Cognitive Skills programme found no significant differences the outcomes of their randomly assigned experimental and comparison groups. Hollin et al. (2004, 2008) and Palmer et al. (2007) reported similar findings in their evaluation of offending behaviour programmes within England and Wales. Likewise, Cann et al. (2003) found no differences between the reconvictions rates of men who had commenced a cognitive skills programme within the UK prison system and a matched control group.

In contrast, the 'treatment received' analysis breaks the experimental group into the naturally occurring groups of programme completers and dropouts and analyses the resulting outcomes of each group separately. In adopting this approach, not only is it

possible to establish whether treatment completion has a positive effects on outcome (what may be provisionally called a 'completion effect'), the researcher can also study the outcomes of the programme dropouts to determine any potential impact of programme non-completion. Further, researchers may also be interested in the impact of different types of programme dropout, for instance, pre-programme (non-starting) and in-programme dropout (non-completion) relating to the comparison and programme completion groups. Finally, such methodology would permit examining the naturally occurring groups to assess for differences between these groups which might be useful operationally in increasing compliance, for instance.

On evaluating the outcomes of their programme completers, dropouts, and the comparison group, van Voorhis et al. (2004) reported that significant differences between the groups emerged. Programme completers were found to have reduced recidivism and time to recidivism compared with the comparison and dropout groups. The comparison and dropout groups also differed, with the dropout group returning the least favourable outcomes. Likewise, Hollin et al. (2008) concluded of their research: "Overall, the evidence suggests that compared to no-treatment controls, there is a positive effect on reconviction of program completion but a higher rate of reconviction for program noncompleters." (p. 281). Cann et al. (2003), Palmer et al. (2007) and Hollin et al. (2004) also reported similar findings; on separating out the experimental group into the naturally occurring groups of completers, non-completers, and non-starters (if possible) differences between the these groups and the comparison group emerged.

One of the problems with the 'treatment received' analysis, however, is that, whether using randomised, matched or designs with statistical control, the comparison of the naturally occurring groups with the full comparison group becomes less justifiable (Gondolf, 2002; Seager, Jellicoe & Dhaliwal, 2004). The comparison group is selected for comparison with the full experimental group: if sub-sections of the experimental group are created through the natural process of attrition, it can be argued that these should be compared with appropriate subsections of the comparison group (Seager, Jellicoe & Dhaliwal, 2004). Indeed, as is reported in Chapter one, completers, non-completers and non-starters are known to systematically differ on a number of variables and a large proportion of these variables are also known to be associated with reconviction outcomes. Hence Debidin and Lovbakke (2005) state:

> The findings [of 'treatment received' design research] may be interpreted as selection effects, that is, that the programme simply served to sort those who would do well anyway from those who would not, regardless of the treatment. The difficulty lies in the inability of the analysis (and the research design) to demonstrate the 'counterfactual', that is, what would have happened in the absence of the programme to the offenders who completed or dropped out? (p. 48)

One potential answer to Seager, Jellicoe and Dhaliwal's and Debidin and Lovbakke's above concerns is to statistically control for differences between the naturally occurring groups. This methodology was adopted by van Voorhis, et al. (2004) and within the Pathfinder papers (Hollin et al., 2004, 2008; McGuire et al., 2008; Palmer et al., 2007). An alternative, however, is to use one-to-one matching of the comparison and experimental groups (Hatcher et al., 2008). By matching each member of the experimental group with a member of the comparison group on the basis of variables that are known to be related to reconviction, it is then possible to compare the completers, non-completers and non-starters to sub-sections of the comparison group. The comparison sub-group members, thus, would represent the 'counterfactual' in that they will have received no intervention. Whilst such evaluation methodology has been utilised in the UK in relation to offence specific programmes (Hatcher et al., 2008), it is yet to be undertaken in relation to the general offending behaviour programmes.

The purpose of this chapter, therefore, is to build on previous outcome evaluation findings utilising a comparison group which has been matched on a one to one basis with the experimental group on key criminogenic factors. This research will evaluate the community delivery of two general offending behaviour programmes, Think First and Enhanced Thinking Skills. Initially, an 'intention to treat' design will assess for differences in reconviction outcomes between an experimental and the matched comparison group. This analysis will be followed by an investigation of the naturally occurring groups of completers, non-completers, and non-starters to determine their reconviction outcomes. These will be assessed alongside those of the full comparison group and then, in line with the thinking of Seager, Jellicoe and Dhaliwal

(2004), in relation to their matched subsections. This chapter will therefore form a basis for the further research presented in this thesis which will investigate the differences between completers, non-completers and non-starters in an effort to advance research in this area. The implications of the results in relation to programme evaluation and practice development will be discussed.

Hypotheses

1. After controlling for time at risk, the group (experimental vs. comparison) variable will significantly predict reconviction outcome.
2. After controlling for time at risk, the group (completer vs. non-completer vs. non-starter vs. comparison) variable will significantly predict reconviction outcome. The resulting odds ratios will indicate that the completers are less likely to be reconvicted that the comparison, non-completers, and non-starter groups.
3. After controlling for time at risk, the group (completer vs. matched comparisons) variable will significantly predict reconviction outcome. The odds ratios will indicate that the completers are less likely that their matched comparisons to be reconvicted.
4. After controlling for time at risk, the group (non-completer vs. matched comparisons) variable will significantly predict reconviction outcome. The odds ratios will indicate that the non-completers are more likely than their matched comparisons to be reconvicted.

The data will also be explored to determine whether collapsed categories (e.g. starters vs. non-starters, drop-outs vs. completers) are able to significantly predict reconviction outcomes after controlling for key variables which differ significantly between the groups (to be determined within univariate analyses).

Method

Procedure and design

The collection of the data for this study was organised and partly undertaken by the researcher as part of a larger study into the effectiveness of offending behaviour programmes (the offending behaviour programmes Pathfinder project). The original dataset relating to the larger study was collected from Probation Areas within England and Wales and represented a national sample of offenders on community rehabilitation orders with the additional requirement to attend an offending behaviour programme. For the purpose of the research to be reported here, a subset of this dataset was utilised. The main Pathfinder research was concerned, in the main, with the reconviction outcomes of programme completers, non-completers, and non-starters comparative to a comparison group of offenders who were under community supervision but had not been ordered to attend an accredited programme. As such basic information (age, risk of reconviction, number of previous convictions, programme completion status, reconviction status) was collected on 2186 offenders sentenced to attend one of three general offending behaviour programmes (Think First, Reasoning and Rehabilitation and Enhanced Thinking Skills). More extensive data were collected, however, on smaller subsets of the overall sample. All those individuals within the final dataset with an order to attend an accredited programme for whom a full set of data (programme information, offender need, demographics, programme suitability, criminal history) were available were therefore selected for the research presented within this and the subsequent chapter (N = 173). Permission to use the data for the purposes of the researcher's PhD submission was granted by the funding body of the original research project, the Home Office Research Development and Statistics department.

As such, this study was of a quasi-experimental design and involved the collection of information relating to two groups of convicted offenders. The first, or 'experimental', group comprised the 173 male offenders described above. All of these offenders had been sentenced to a community rehabilitation order or a community

punishment and rehabilitation order[6] in 2002. All had a requirement attached to their order to attend a general offending behaviour programme, either the Enhanced Thinking Skills programme or the Think First programme, at their local probation service office. Initial data were gathered from databases and files held by the Probation Areas responsible for delivering the programmes. The data collected were the offenders' name, date of birth, date of sentence, and allocated programme.

The second group comprised a pool of 2749 male individuals from which a comparison group was selected through one to one matching with the experimental group. The data collected in relation to this pool derived from two sources: the Probation Index, a national database of those individuals on a community sentence, and files held by probation areas. All individuals within this pool had been sentenced to a community rehabilitation order but this time without the requirement to attend an offending behaviour programme of any kind. The data collected for this larger pool comprised the offenders' name, date of birth, and date of sentence.

Participants

As outlined above, the experimental group was comprised of a sample of 173 sentenced offenders with a community rehabilitation order[7] with a requirement to attend a general offending behaviour programme (either Think First or Enhanced Thinking Skills). An additional 173 participants were selected, using the method to be described below, from the larger pool of 2749 offenders who had received a community sentence with no requirement to attend an offending behaviour programme to form a comparison group. The average age of those in the study was 26.85 (*SD* = 7.09) with a range of ages from 18 to 55 years. The mean Offender Group Reconviction Score

[6] Once the full dataset was purged of those for whom full data were not available, all those offenders on the Reasoning and Rehabilitation programme were lost.

[7] Under the Criminal Justice Act 2003 (and from April 2005) these orders now have the generic term of 'Community Orders'. An accredited programme can be undertaken as a requirement of the community order. At the time of data collection, however, accredited programmes were delivered as part of either a community rehabilitation order or a community punishment and rehabilitation order.

(OGRS2)[8] score was 62.04 (*SD* = 21.60) with a range scores from 8 to 98 and the mean number of previous convictions[9] was 8.60 (*SD* = 8.36) with a range from 0 to 55.

Measures

Reconviction data on the individuals within both the experimental group and the comparison group pool were retrieved from the Offenders Index (OI). The OI is a national database of conviction information for all offenders which is maintained centrally. The data retrieved from this source were details of any reconvictions (offence type, date of reconviction and so on) and the OGRS2 risk score. This latter variable, along with offender age and the number of previous convictions, was used in the one to one matching of the experimental group to the comparison group.

Ethical considerations

As this research was based on a file review of information held within Probation Areas, the participants were not individually involved in the research process. At the time of data collection, the majority of participants had completed their probation orders and were no longer in contact with the Probation Area. As such, it would have been intrusive to the participants of the research to be contacted to request their informed consent for their data to be used within the research. Additionally, the data was deemed to be under the ownership of each local Probation Area rather than the individual offenders. As such, informed consent was not gained from the participants. Consent to use the data for the purpose of research was instead gained from each local Probation Area and from the National Probation Directorate.

The data were stored electronically within secure files. The only identifying information retained within the data files were those required to match the dataset to the OI. Once this matching process was complete, all identifying information was deleted from the files. As such, from analysis of the final dataset it is not possible to identify to which individuals the data refer.

[8] The OGRS2 risk score is an estimate of an individual offender's probability of reconviction within two years based on nine demographic and criminal history variables score (Taylor, 1999).
[9] The 'previous convictions' variable relates to the number of court appearances at which the individual has been reconvicted.

Matching strategy

To ensure that the comparison group resembled the experimental group as closely as possible, one to one matching was undertaken. This procedure reduced the pool of 2749 down to a smaller group of 173 individuals matched on a one to one basis according to OGRS2 score, number of previous convictions, and the age of the participant. The matching process was completed manually by matching each individual one at a time. The procedure was conducted in three steps: first, the experimental group individual's OGRS2 score was matched against the experimental group pool to produce a sub-sample of those who had the same risk score. In all cases, there was at least one member who matched the risk level exactly; indeed in most cases there were more than one. Second, from this OGRS2 matched sub-sample a smaller sub-group was matched to the experimental group individual according to their age. Where there were no individuals within the sub-sample with precisely the same age, the closest possible match was selected. Where there was only one closest possible match, this individual was selected regardless of their number of previous convictions. Where there was more than one closest possible match (in either direction: older or younger), these were selected and carried on through to the third step in the process. Third, the procedure was repeated, this time matching on the number of previous convictions each individual had. Again, if this was not possible the closest possible matches (again in either direction) were taken. If more than one match (or closest possible match) remained at the end of this process, the selection of the comparison group individual was made randomly from the matched pool members; the participant identification numbers (which were present for all individuals within the file) were written on pieces of paper, folded, placed on the desk and mixed up. The researcher blindly selected one of these pieces of paper and selected the individual corresponding to the identification number written on it.

As a consequence, all resulting pairs had the same OGRS2 score and the closest available matches in the number of previous convictions and age. The mean OGRS2 scores, age, and number of previous convictions are displayed in table 2.1. Given the matching procedures employed, these values are highly similar across groups.

An independent t test comparing the groups' OGRS2[10] scores was non-significant (*t* [344] = 0.00, *p* = 1, Cohen's *d* = 0). As the age and number of previous convictions variable distributions (see Appendix A for the normality curves) were non-normal (age: comparison group, Kolmogorov-Smirnov Z (*K-S Z*) = 1.488, *p* = 0.024, experimental group, *K-S Z* = 1.594, *p* = 0.012; number of previous convictions: comparison group, *K-S Z* = 2.272, *p* < 0.001, experimental group, *K-S Z* = 2.228, *p* < 0.001) Mann Whitney U tests were employed on these data. As expected, these were both non-significant (age: *U* = 14673, *p* = 0.754, *r* = 0.01; number of previous convictions: *U* = 14925, *p* = 0.966, *r* = 0.002).

Table 2.1. Descriptive Statistics by Treatment Allocation Group

	Experimental (*N* = 173)	Comparison (*N*= 173)
OGRS2 score	62.04 [a] (21.63 [b])	62.04 (21.63)
Age	26.92 (7.27)	26.57 (6.92)
Number of previous convictions	8.66 (8.53)	8.53 (8.21)

[a] Mean [b] Standard Deviation

Analysis strategy

This evaluation aimed to determine the impact of offending behaviour interventions on subsequent reconviction rates. It was necessary, however, to make adjustments for the length of time that each individual had the opportunity to reconvict.

[10] Given that the OGRS2 score is calculated using a statistical algorithm (see Appendix B) it was necessary to consider whether the data were appropriate for use within parametric analyses. The OGRS2 algorithm produces a score from 0-100 which represents an estimate of the expected probability of a group of offenders who match that individual on the set of OGRS2 variables being reconvicted within two years (Copas & Marshall, 1998, p. 170). As such, OGRS2 data meet the parametric assumption of being at least interval (Field, 2005); equal intervals on the scale represent equal differences as each interval of one represents one percent. Indeed, it could said that OGRS2 data is actually ratio data (in that an offender with a score of 60 comes from a group who is twice as likely to be reconvicted within two years as an offender with a score of 30) and hence, as such, is even more suitable for parametric analyses. The data also meet the assumption of independence: the score of one offender does not influence the score of another. Finally, the assumptions of normality and homogeneity of variance are assessed within this thesis for each sample utilised. Where these assumptions are violated for that sample non-parametric analyses are used. Where they are not, given the above conclusions in relation to the nature of the data, parametric analyses are used.

As the Offenders Index reconviction information was collected on a fixed census date, the obtained data includes details of any reconviction from the individuals' sentence dates up to this census date. With the sentence dates naturally varying between individuals, the time that the individual is 'at risk' of reconviction(s) also naturally varies. Across the sample, the time at risk ranged from 280 to 1454 days, with a mean of 760.09 (*SD* = 346.83). Statistical control of the 'time at risk' variable (the number of days between sentence date and the census date) was therefore employed.

The analysis first compared the experimental group to the matched comparison group using a design which is analogous to an 'intention to treat' design. This analysis provides an evaluation of the policy rather than an evaluation of treatment as it includes all those who failed to start or complete a programme within the experimental group. Following this, a 'treatment received' analysis compares the reconviction rates of the naturally occurring groups (programme completers, programme non-completers and programme starters) to their matched comparisons and each other. This provides an analysis of the impact of the intervention rather than the policy as it allows a comparison of those who received the full dosage of treatment against matched controls and those who failed to start or complete the programme.

Where possible, the effect size correlation or Φ (phi) statistic was calculated to determine the magnitude of relationships between the variables (Kotrlik & Williams, 2003). The effect size correlation was calculated from the single degree of freedom χ^2 value using the formula $\Phi = \sqrt{(\chi^2 (1) / N)}$. For contingency tables greater than a two by two design, Cramer's Φ was instead calculated (Kotrlik & Williams, 2003) using the formula $\Phi_c = \sqrt{(\chi^2 / N (k\text{-}1))}$, where k is the smallest number of levels between the variables. For Mann Whitney analyses, the effect size correlation was calculated using the formula $r = Z / \sqrt{(\text{number of observations})}$ (Field, 2005), and for ANOVA tests, the partial eta squared statistic (η^2) has been taken from the SPSS output and square rooted to obtain *r*. For independent sample t-tests Cohen's *d* was calculated from the difference of the sample means/pooled standard deviation.

Results

Treatment allocated (experimental vs. comparison group) analysis:

Univariate analysis: At the census point, 227 of the 346 offenders (65.61%) had been reconvicted. A total of 125 (72.3%) of the comparison group were reconvicted compared to 102 (59%) of the experimental group. This difference was statistically significant χ^2 (1, N = 346) = 6.78, p = 0.009, Φ = 0.14). The average time at risk scores for the comparison group were, however, longer than those of the experimental group (1019.78 and 500.40 respectively) and so the univariate difference should be interpreted with caution.

Multivariate analysis: A sequential logistic regression analysis was performed on the data to determine the predictive ability of group membership (experimental or comparison) on reconviction after the effects of time at risk were controlled for. At block one, the time at risk variable was entered into the analysis, followed by the group variable at block two.

At block one, the time at risk variable did not produce a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, N = 346) = 19.62, p = 0.012. This model was, however, significantly better than the constant-only model (block 0) containing only the intercept but no predictor variables: χ^2 (1, N = 346) = 18.63, p < 0.001. The addition of group (experimental, comparison) at stage two did not lead to a significant improvement in the model provided by the time at risk variable, χ^2 (1, N = 346) = 0.778, p = 0.378. The goodness of fit of the model as measured by the Hosmer and Lemeshow Test was not good, χ^2 (8, N = 346) = 21.92, p = 0.005, with 65.3% correct classification of cases. Table 2.2 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. The odds ratio (Exp [B]) indicates that for comparison group members, the odds of being reconvicted are 0.744 times as large as the odds of an experimental group member being convicted. Put more simply for every three comparison group members reconvicted, there will be four experimental group members also reconvicted. Despite this finding, the addition of group into the model was not statistically significant.

Table 2.2: Logistic regression of reconviction as a function of time at risk and treatment group (experimental, comparison)

	B	SE	Wald	Exp (B)	95% CI for Exp (B)
Time at risk	0.002	0.001	12.219****	1.002	1.001 – 1.003
Group	-0.296	0.335	0.779	0.744	0.385-1.435
Constant	-0.548	0.302	3.289	5.78	

**** $p < 0.001$

Treatment received analysis (analysis of naturally occurring groups)

One of the issues with the above analysis is that any completion effect is at risk of being diluted by combining the programme completers with those who failed to complete or even start the programme. In combining these individuals within one experimental group, it is not possible to determine the effect of the full dosage of the programme on those who completed it. The following analysis, therefore, aims to address this limitation by analysing the naturally occurring groups of programme completers, non-completers, and non-starters separately.

Univariate analysis: Of the 173 offenders in the experimental group, 66 (38.2%) completed the programme (Completers), 38 (22.0%) started the programme but failed to complete it (Non-completers) and 69 (39.9%) failed to start the programme at all (Non-starters)[11]. Of these groups 26 (39.4%) of the completers, 26 (68.4%) of the non-completers, and 50 (72.5%) of the non-starters were reconvicted. Within the comparison group, 125 of the 173 (72.3%) were reconvicted. A chi-square analysis of group by reconviction indicated a significant association (χ^2 (3, N = 346) = 25.057, $p <$ 0.001, $\Phi_c = 0.27$). This univariate difference should, however, be interpreted with caution as there has been no control for time at risk (or indeed any other variables that might differ systematically between the groups) within this analysis.

[11] Given the small sample size within the non-completer sample, it could be argued that this analysis is underpowered. As such conclusions drawn should rely not only on the p-values but also on the odds ratios and effect sizes and should be considered tentative.

The descriptive information for the four groups in terms of OGRS2 score, age, and number of previous convictions is shown in Table 2.3.

Table 2.3. Descriptive Statistics by Treatment Received Groups

	Completers (*N* = 66)	Non-Completers (*N* = 38)	Non-Starters (*N* = 69)	Comparison (*N* = 173)
OGRS2 score	56.38 [a] (22.47 [b])	63.84 (19.22)	66.46 (21.13)	62.04 (21.63)
Age	28.38 (7.80)	24.66 (5.64)	26.77 (7.31)	26.57 (6.92)
Previous convictions	8.05 (7.13)	7.76 (7.95)	9.75 (9.95)	8.53 (8.21)
Time at risk	406.83 (87.78)	562.63 (97.00)	555.62 (96.55)	1019.78 (302.59)

[a] Mean [b] Standard Deviation

Kolmogorov-Smirnov tests indicated that the parametric assumption of normality was violated for at least one group for the Age (comparison group: *K-S Z* = 1.488, *p* = 0.024), number of previous convictions (comparison group: *K-S Z* = 2.228, *p* < 0.001; completers: *K-S Z* = 1.655, *p* = 0.008; non-starters: *K-S Z* = 1.484, *p* = 0.024) and time at risk (comparisons: *K-S Z* = 1.828, *p* = 0.003). Comparisons between the groups on these variables were therefore determined using the non-parametric Kruskall-Wallis test. There were no differences between the groups' age (χ^2 (3, *N* = 346) = 5.79, *p* = 0.122, Φ_c = 0.09) and number of previous convictions (χ^2 (3, *N* = 346) = 1.71, *p* = 0.635, Φ_c = 0.05). The time at risk variable did, however, vary significantly between groups: χ^2 (3, *N* = 346) = 202.78, *p* < 0.001, Φ_c = 0.54. Pairwise comparisons were all statistically significant, bar the non-completer and non-starter comparison (comparisons vs completers: *U* = 373, *p* < 0.001, *r* = 0.72; comparisons vs non-completers: *U* = 787, *p* < 0.001, *r* = 0.50; comparisons vs non-starters: *U* = 1380, *p* < 0.001, *r* = 0.60; completers vs non-completers: *U* = 329, *p* < 0.001, *r* =0.61; comparisons vs non-starters: *U* = 642.50, *p* < 0.001, *r* =0.62; non-completers vs non-starters: *U* = 1266, *p* = 0.770, *r* =0.03). A one way ANOVA found no significant differences between the groups' OGRS2 scores (F = 2.601, *df* = 3, 345, *p* = 0.052, *r* = 0.47).

Multivariate analysis: A sequential logistic regression was next performed to assess the impact of the four groups on reconviction status. To control for the effects of time at risk, which was found to differ between groups, this variable was entered at block one with the group variable entered at block two.

At block one, and as with the treatment allocated design, the time at risk variable did not produce a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, *N* = 346) = 19.62, *p* = 0.012 but the model was significantly better than the constant-only model: χ^2 (1, *N* = 346) = 18.63, *p* < 0.001. The addition of group (comparison, completers, non-completers, non-starters) at stage two led to a significant improvement in the model provided by the time at risk variable, χ^2 (3, *N* = 346) = 12.291, *p* = 0.006. The fit of the model as measured by the Hosmer and Lemeshow Test was good, χ^2 (8, *N* = 346) = 12.288, *p* = 0.139, with 69.7% correct classification of cases. Table 2.4 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables.

Table 2.4: Logistic regression of reconviction as a function of time at risk and treatment group (comparison, completer, non-completer, non-starter)

	B	SE	Wald	Exp (B)	95% CI for Exp (B)
Time at risk	0.001	0.001	6.903**	1.001	1.000 – 1.002
Group			11.709**		
Group (1)	-0.628	0.391	2.582	0.534	0.248 – 1.148
Group (2)	-1.195	0.377	10.044***	0.303	0.145 – 0.634
Group (3)	-0.205	0.442	0.215	0.815	0.343 – 1.937
Constant	0.192	0.399	0.232	1.212	

** *p* < 0.01, *** *p* < 0.005

The logistic regression was repeated but this time with pairs of groups entered at block two rather than all four groups. Table 2.5 displays the odds ratios for each comparison. In each case, the group in the left column was the reference category.

Table 2.5: Odds ratio for each pairwise comparison

	Completers	Non-completers	Non-starters
Comparisons	1.827	0.653	0.528
Completers	-	0.348	0.299*
Non-completers	-	-	0.808

* $p < 0.05$

The odds ratios displayed here indicate that comparison group members are 1.827 times as likely to be reconvicted as the completer group, or more simply, for every nine comparison group members reconvicted, there will be approximately 5 completer group members reconvicted.

Odds ratios of less than one for all other comparisons indicates that those in the rows (the reference category) perform better in terms of reconviction than those in the columns: comparison group members are less likely to be reconvicted than the non-completers and the non-starters, completer group members are less likely to be reconvicted than the non-completers and non-starters and non-completer group members are marginally less likely than the non-starters to be reconvicted.

Analysis of matched naturally occurring groups in contrast to their matched comparisons:

The above analysis served to compare the four groups (comparison, completers, non-completers and non-starters) in relation to their reconviction outcomes. With a comparison group which is matched on a one to one basis, however, it would make sense to compare each of the naturally occurring groups with the portion of the comparison group with which those individuals have been matched. The three analyses that follow, therefore, repeat the above analysis with each of the naturally occurring groups (completers, non-completers and non-starters) and their matched comparison pairs. As these groups have been matched on OGRS2, age, and number of previous convictions, it was only necessary to control for the time at risk differences between the groups.

Completer and matched comparison: The time at risk variable was entered into a logistic regression at block one and the group (completer vs. matched comparison) variable subsequently entered at block two. At block one, the time at risk variable produced a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, *N* = 132) = 9.93, *p* = 0.270. This model was significantly better than the constant-only model (block 0) containing only the intercept but no predictor variables: χ^2 (1, *N* = 132) = 12.32, *p* < 0.001. The addition of group (completer, matched comparison) at stage two did not lead to a significant improvement in the model provided by the control variables, χ^2 (1, *N* = 132) = 0.886, *p* = 0.347. The goodness of fit of the model as measured by the Hosmer and Lemeshow Test was not good, χ^2 (8, *N* = 132) = 18.45, *p* = 0.02, with 64.4% correct classification of cases. Table 2.6 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. The odds ratio indicates that for the comparison group, the odds of being reconvicted are 1.717 as large as the odds of a completer being reconvicted; for every three completers reconvicted, there will be five comparison group members reconvicted.

Table 2.6: Logistic regression of reconviction as a function of time at risk and treatment group (completer, matched comparison)

	B	SE	Wald	Exp (B)	95% CI for Exp (B)
Time at risk	0.001	0.001	2.024	1.001	1.000 – 1.003
Group	0.540	0.576	0.879	1.717	0.555 – 5.313
Constant	-0.886	0.408	3.717	0.412	

Non-completer and matched comparison analysis: This analysis was repeated for the non-completers and their matched comparisons. At block one, the time at risk variable produced a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, *N* = 76) = 6.08, *p* = 0.638. This model was not significantly better than the constant-only model (block 0): χ^2 (1, *N* = 76) = 0.705, *p* = 0.401. The addition of group (non-completer, matched comparison) at stage two did not lead to a significant improvement

in the model provided by the control variables, χ^2 (1, *N* = 76) = 0.777, *p* = 0.378. The goodness of fit of the model as measured by the Hosmer and Lemeshow Test was good, χ^2 (8, *N* = 76) = 12.51, *p* = 0.130, with 67.1% correct classification of cases. Table 2.7 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. The odds ratio of 0.536, however, indicates that the odds of a comparison group member being reconvicted are 0.536 as large as the odds of a non-completer being convicted; for every one comparison group member reconvicted, there will be approximately two non-completers reconvicted.

Table 2.7: Logistic regression of reconviction as a function of time at risk and treatment group (non-completer, matched comparison)

	B	SE	Wald	Exp (B)	95% CI for Exp (B)
Time at risk	0.001	0.001	1.456	1.001	0.999 – 1.004
Group	-0.623	0.706	0.780	0.536	0.135 – 2.138
Constant	0.007	0.723	0.993	1.007	

Non-starter and matched comparison analysis: This analysis was repeated, once more, for the non-starters and their matched comparisons. At block one, the time at risk variable produced a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, *N* = 138) = 7.59, *p* = 0.475. This model was not significantly better than the constant-only model (block 0), although it did approach significance: χ^2 (1, *N* = 138) = 3.746, *p* = 0.053. The addition of group (non-completer, matched comparison) at stage two did not lead to a significant improvement in the model provided by the control variables, χ^2 (1, *N* = 138) = 0.763, *p* = 0.382. The goodness of fit of the model as measured by the Hosmer and Lemeshow Test was good, χ^2 (8, *N* = 138) = 3.57, *p* = 0.894, with 75.4% correct classification of cases. Table 2.8 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. The odds ratio indicates that the odds of comparison group members are 0.616 as large as the odds of the non-starters being reconvicted; for every three comparison group members reconvicted, there will be approximately five non-starters reconvicted.

Table 2.8: Logistic regression of reconviction as a function of time at risk and treatment group (non-completer, matched comparison)

	B	SE	Wald	Exp (B)	95% CI for Exp (B)
Time at risk	0.002	0.001	3.796	1.002	1.000 – 1.004
Group	-0.485	0.551	0.773	0.616	0.209 – 1.814
Constant	-0.067	0.593	0.013	0.935	

Analysis of starters, non-starters, completers and drop outs:

A final type of analysis was undertaken with these data. Within this analysis, the naturally occurring groups were collapsed. This allowed the comparison, first, of all those who started a programme (starters: comprised of non-completers and completers) with all those who did not start a programme (non-starters) and second, of those who failed to complete a programme (dropouts: comprised of non-completers and non-starters) with those who did complete (completers).

Starters and Non-starters: Table 2.9 shows the mean age, OGRS2, Number of previous convictions and Time at risk of the Starters and Non-starters.

Table 2.9: Descriptive Statistics by Treatment by Starter and Non-starter groups

	Starters (*N* = 104)	Non-starters (N = 69)
OGRS2 score	59.11 [a] (21.55 [b])	66.46 (21.13)
Age (years)	27.02 (7.28)	26.77(7.31)
Number of previous convictions	7.94 (7.40)	9.75 (9.95)
Time at risk	463.76 (118.00)	555.62 (96.55)

[a] Mean [b] Standard Deviation

As these groups are obviously not matched, analyses were conducted to assess for significant differences between the groups on the variables of age, OGRS2, previous

convictions, and time at risk. Non-normal distributions were found for the Starters' and Non-starters' previous convictions values (*K-S Z* = 1.94, *p* = 0.001 and *K-S Z* = 1.48, *p* = 0.024 respectively) hence this variable was analysed using the Mann Whitney non-parametric test. This analysis returned non-significant results: *U* = 3196.00, *p* = 0.223, *r* = 0.09. Independent sample t-tests between the starters and non-starters indicated significant differences between the groups' OGRS2 scores (*t* [173] = -2.216, *p* = 0.028, *d* = 0.344), and time at risk (*t* [173] = -5.38, *p* < 0.001, *d* = 0.852). The age variable did not differ significantly between the groups (*t* [173] = 0.222, *p* = .825, *d* = 0.034).

The two groups' reconviction statistics are displayed in table 2.10. The frequency of reconviction across the two groups, unadjusted for between-group variations, differed significantly from chance, χ^2 (1, *N* = 173) = 8.650, *p* = 0.003, Φ = 0.22. The lower rate of reconviction was in the Starters group.

Table 2.10: Reconviction by group

	Not-reconvicted (%)	Reconvicted (%)	Total (%)
Starters	52 (50.00)	52 (50.00)	104 (100)
Non-starters	19 (27.50)	50 (72.50)	60 (100)
Total	71 (41.00)	102 (59.00)	173 (100)

A sequential logistic regression was carried out to examine the effect of treatment groups in the prediction of reconviction. The first stage of the logistic regression gives a model showing how the control variables – OGRS2 scores and Time at risk – are related to the outcome variable (reconviction). At the second stage, the group variable was added to the model. The intercorrelations between the control variables were examined. They did not exceed the level (>0.70) that would indicate multicollinearity (Tabachnick & Fidell, 2006).

The control variables at block one produced a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, *N* = 173) = 8.205, *p* = 0.414. This model was significantly better than a constant-only model containing only the intercept but no predictor variables: χ^2 (2, *N* = 173) = 14.493, *p* = 0.001. The addition of group (Starter,

Non-starter) at block two led to a non-significant improvement in the model provided by the control variables, $\chi^2(1, N = 173) = 3.072$, $p = 0.080$. The goodness of fit as measured by the Hosmer and Lemeshow Test was, however, good, $\chi^2 (8, N = 173) = 10.074$, $p = 0.260$, with 65.3% correct classification of cases. Table 2.11 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. The odds ratio indicates that the odds of a programme starter being reconvicted are 0.533 times as large as the odds of a programme non-starter being reconvicted. Again, this means that for each programme starter that is reconvicted, there will be approximately two non-starters reconvicted.

Table 2.11: Logistic regression of reconviction as a function of OGRS2, time at risk, and treatment group (starters, non-starters)

	B	SE	Wald	Exp (B)	95% CI for Exp (B)
OGRS2	1.424	0.776	3.368	4.153	0.908-19.001
Time at risk	0.003	0.002	4.027*	1.003	1.000-1.006
Group	-0.630	0.362	3.036	0.533	0.262-1.082
Constant	-1.612	0.950	2.881	0.199	

* $p < 0.05$

Completers and Dropouts: Additional analyses were conducted to assess whether there were any differences between Completers and Dropouts. The categories of Non-starters and Non-completers were therefore collapsed into the category of Dropouts. Table 2.12 shows the mean age, OGRS2, Number of previous convictions and Time at risk.

Analyses were conducted to assess for significant differences between the groups on the variables of age, OGRS2, previous convictions, and time at risk. The Kolmogorov-Smirnov one-sample test detected non-normal distributions relating to the Completers' and Dropouts' previous convictions values (*K-S Z* = 1.65, *p* = 0.008 and *K-S Z* = 1.72, *p* = 0.005 respectively) and the Dropouts' Age values (*K-S Z* = 1.43, *p* = 0.033). Hence these data were analysed using the Mann Whitney non-parametric test. The

analysis of Previous convictions returned non-significant results: $U = 3395.00$, $p = 0.670$, $r = 0.03$. Age, however, did differ significantly between groups: $U = 2898.50$, $p = 0.048$, $r = 0.15$. Independent sample t-tests indicated significant differences between the groups' OGRS2 scores (t [173] = -2.756, $p = 0.006$, $d = 0.42$), and time at risk (t [173] = -10.375, $p < 0.001$, $d = 1.64$).

Table 2.12: Descriptive Statistics by Treatment by Completer and Dropout groups

	Completers (N = 66)	Dropouts (N = 107)
OGRS2 score	56.38[a] (22.47[b])	65.53 (20.42)
Age (years)	28.38 (7.80)	26.02(6.81)
Number of previous convictions	8.05 (7.13)	9.05 (9.30)
Time at risk	406.83 (87.77)	558.11 (96.31)

[a] Mean [b] Standard Deviation

The two group's reconviction data are displayed in table 2.13. The frequency of reconviction across the two groups, unadjusted for between-group variations, differed significantly from chance, χ^2 (1, N = 173) = 16.992, $p < 0.001$, $\Phi = 0.31$. The lowest rate of reconviction was in the Completers group.

Table 2.13: Reconviction by group

	Not-reconvicted (%)	Reconvicted (%)	Total (%)
Completers	40 (60.61)	26 (39.39)	66 (100)
Dropouts	31 (28.97)	76 (71.03)	107 (100)
Total	71 (41.04)	102 (58.96)	173 (100)

A sequential logistic regression was performed to examine the effect of treatment groups in the prediction of reconviction. The first stage of the logistic regression gives a model showing how the control variables – Age, OGRS2 scores and

Time at risk – are related to the outcome variable (reconviction). At the second stage, the two groups are added to the model given by the control variables. As multicollinearity between predictor variables can cause problems in multivariate analysis, the intercorrelations between the control variables were examined. The magnitude of these correlations did not exceed the level (>0.70) that would point to multicollinearity (Tabachnick & Fidell, 2006).

First, the control variables were entered into the analysis and, as earlier, they produced a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, *N* = 173) = 5.820, *p* = 0.667. This model was significantly better than a constant-only model containing only the intercept but no predictor variables: χ^2 (2, *N* = 173) = 21.768, *p* < 0.001. Second, the addition of group (Completer, Dropout) led to a significant improvement in the model provided by the control variables, χ^2 (1, *N* = 173) = 4.181, *p* = 0.041. The goodness of fit of the model as measured by the Hosmer and Lemeshow Test was good, χ^2 (8, *N* = 173) = 12.611, *p* = 0.126, with 65.3% correct classification of cases. Table 2.14 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. The odds ratio indicates that the odds of a programme completer being reconvicted are 0.417 times as large as the odds of the dropouts being reconvicted; for two completers reconvicted there will be approximately five dropouts reconvicted.

Table 2.14: Logistic regression of reconviction as a function of age, OGRS2, time at risk, and treatment group (completers, dropouts)

	B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Age	-0.054	0.024	4.973*	0.947	0.904-0.993
OGRS2	1.100	0.799	1.894	3.004	0.627-14.390
Time at risk	0.002	0.002	1.191	1.002	0.998-1.006
Group	-0.875	0.430	4.136*	0.417	0.179-0.969
Constant	0.511	1.227	0.174	1.667	

* p < 0.05

Review of the hypotheses

Given the number and complexity of the analyses undertaken, a review of how the analysis relates to the hypotheses is warranted.

1. After controlling for time at risk, the group (experimental vs. comparison) variable will significantly predict reconviction outcome.

   This hypothesis was not supported. The logistic regression did not significantly predict group outcome from the group variable after controlling for time at risk. This should be interpreted with caution, however, as the fit of the data to the model was not good.

2. After controlling for time at risk, the group (completer vs. non-completer vs. non-starter vs. comparison) variable will significantly predict reconviction outcome. The resulting odds ratios will indicate that the completers are less likely to be reconvicted than the comparison, non-completers, and non-starter groups.

   Both components of this hypothesis were supported by the analysis. The logistic regression should be interpreted with caution, however, as the fit of the data to the model was not good.

3. After controlling for time at risk, the group (completer vs. matched comparisons) variable will significantly predict reconviction outcome. The odds ratios will indicate that the completers are less likely that their matched comparisons to be reconvicted.

   The group variable (completer vs. matched comparison) did not significantly predict reconviction outcomes. However, the odds ratio did indicate that for every three completer reconvicted there would be five comparison group members.

4. After controlling for time at risk, the group (non-completer vs. matched comparisons) variable will significantly predict reconviction outcome. The odds ratios will indicate that the non-completers are more likely than their matched comparisons to be reconvicted.

This hypothesis was not supported. However, the odds ratios did indicate that the non-completers were more likely than their matched comparisons to be reconvicted. This should be interpreted with caution, however, as it was not statistically significant.

Discussion

This study aimed to determine the effect of participation in one of two general offending behaviour programmes on reconviction outcomes within the available sample. As such, this study aimed to build on previous research findings (McMurran & Theodosi, 2007; Hollin et al., 2008; McGuire et al., 2008) by utilising a quasi-experimental design which matched the experimental and comparison groups on a one to one basis. Such a design permitted the assessment of the full experimental group relative to the matched comparison group and the assessment of the naturally occurring completers, non-completers and non-starters relative to their individually matched comparisons. The analyses presented here demonstrate that, despite suffering from a lack of power, the pattern of results obtained was similar to those of previous research utilising different methodologies.

As found by Hollin et al. (2008) and McGuire et al. (2008), a 'treatment allocated' analysis comparing the experimental and matched comparison groups' outcomes found there to be no impact of group membership on reconviction. Analysis of the odds ratios in the present study, however, revealed the comparison group were less likely to be reconvicted than the experimental group. Such a finding might initially lead to a questioning of the efficacy of offending behaviour programmes; with the experimental group performing less favourably than the comparison group it might be concluded that there is a damaging effect of programme participation. This finding, however, needs to be considered in context.

In accord with the research of Hollin et al. (2008) and McGuire et al. (2008), the rate of programme completion amongst the current sample was low; only 66 of the sample of 173 offenders sentenced to attend a programme completed the full programme. Hence almost two thirds of the experimental group failed to receive the full dosage of a programme which is designed, in line with accreditation criteria, to be of a

length and duration appropriate to the needs of the targeted population. As such, it is likely that, in a 'treatment allocated' design, any potential effect of programme completion is diluted by the effect of programme dropout (Hollin et al., 2008; McGuire, et al., 2008; van Voorhis, et al., 2004).

It would seem, however, that this explanation is not sufficient to describe the findings reported here and within other commentators' work. If a positive effect of treatment on the completers was weakened by a null effect on the programme dropouts, we would still expect the experimental group to perform slightly better, or at worse equal to, the comparison group. The results reported here, however, find the comparison group with lower reconviction rates than the experimental group. Meta-analytic research by McMurran and Theodosi (2007) has provided evidence that may explain such findings. Following their review of treatment outcomes amongst non-completers, these authors concluded that programme dropout may not merely result in a lack of an effect but that it may actually be detrimental to the sample's subsequent reconviction rates. As such, it is not surprising that any potentially positive completion effects returned by just over a third of the sample might be swamped by negative effects of the larger sample of programme dropouts. Such findings were also reported in the programme evaluations of Hollin et al. (2008) and McGuire et al. (2008). Within these evaluations, the authors recommend that to evaluate the efficacy of an intervention (as opposed to a policy) it is necessary to investigate the completers separately from the non-completers or non-starters of that intervention.

To tease out such effects, therefore, the naturally occurring groups of completers, non-completers, and non-starters were assessed against the full comparison group (the 'treatment received' analysis). In this study, and in line with the findings of Hollin et al. (2008), McGuire et al. (2008), and van Voorhis et al. (2004), completers were less likely than all other groups, including the comparison group, to be reconvicted. Whilst the comparison between completers and non-starters was the only one to reach statistical significance, the odds ratios indicated moderate effects of completion comparative to the comparison, non-completer and non-starter groups. Indeed according to the odds ratios that were obtained for every one programme completer subsequently reconvicted there will be three non-completers and three non-starters also receiving reconvictions. The completers also fare well again the full

comparison group: for every five completers reconvicted, nine comparison group members will also be reconvicted. These findings indicate the benefits of programme completion over programme dropout and over no intervention.

These conclusions, however, could be criticised for being misleading. Seager, Jellicoe and Dhaliwal (2004) have argued that the practice of using the full comparison group to evaluate the outcomes of a completer group is methodologically weak. They propose that failing to remove the proportion of the comparison group that would non-complete (if given the opportunity) but removing the non-completers from the experimental group fails to compare like with like. Whilst the authors advocate the mathematical removal of the effect of non-completers (but do not explain how this should be estimated), the above analysis instead advocated a one-to-one matching process to establish discrete sections of the comparison group which correspond, on a one-to-one basis, with the completers, non-completers and non-starters respectively. The variables upon which the matching was undertaken are those known to be associated with reconviction and hence, all things being equal, it would be expected that the matched groups would have similar reconviction outcomes. As such, any differences observed here could be attributed, at least tentatively to the experience of programme completion. The matched comparison analyses, however, failed to significantly predict reconviction from the group (completer vs. matched comparison) variable. However, the odds ratios indicated that for each three completers reconvicted, five matched comparison group members would also be reconvicted. There is hence a possibility that the unpowered nature of this analysis has resulted in a failure to detect significant prediction.

In relation to the investigation of the effect of non-completion, this research reported mixed findings. The first point to note was the rate of dropout amongst the sample: 62% dropped out of the programme; consisting of 22% who commenced but fail to complete and 40% who failed to attend even at session one. The comparison of these groups with the full matched comparison group revealed their reconviction outcomes, at least according to the odds ratios, to be unfavourable: non-completers and non-starters were more likely to be reconvicted (at a ratio of almost 2:1) than the full matched comparison group. Of particular interest, however, is whether there was evidence within these data of a ‘non-completion effect’. Taking the argument of Seager,

Jellicoe and Dhaliwal (2004), to determine the effect of non-completion the non-completer group should be compared to the sub-group of the comparison group that would have also dropped out given the chance. When compared against their one-to-one matches, there was a non-significant association between non-completion and reconviction. On the face of it, therefore, it would seem that there is no evidence from this research for the existence of a detrimental 'non-completion effect'. However, the odds ratio indicated that non-completers are twice as likely as their matched comparison group to be reconvicted. Again, there is a possibility that the lack of power within this analysis has resulted in a non-significant prediction. Further research with larger sample sizes should be undertaken to address this possibility.

The implications of these findings are perhaps clouded by the methodological difficulties encountered in conducting intervention evaluations. The 'treatment allocated' design, performed here to replicate previous results, has been shown to lack practical usefulness when attrition from the intervention under evaluation is high and the effect of non-completion is likely to be negative. Indeed, McGuire et al. (2008) have stated that "whereas a direct comparison between experimental and control samples can test hypotheses regarding the effect of offering treatment, only a 'treatment received' analysis that separates treated from untreated samples can allow evaluation of meaningful treatment effects" (McGuire et al., 2008, p.36).

The 'treatment received analysis' which provides tentative evidence of a positive completion effect and a negative non-completion effect, however, is undermined by the failure to control for pre-existing differences between the groups which have the potential to explain the effects reported (Debidin & Lovbakke, 2005; Hollin et al., 2008; McGuire et al., 2008). Within this study, a matching process was undertaken to ensure that the comparison group and experimental groups were comparable; however, it was only possible to match the samples on age, risk of reconviction scores, and the number of previous convictions. As a result it is not possible to rule out the possibility that unmeasured, and hence uncontrolled for, variable(s) are responsible for the split of the sample into completers, non-completers and non-starters and, consequently, the observed effects. Indeed, within this chapter significant pre-programme differences were observed between the programme completers and programme dropouts in relation to age and risk of reconviction scores. Programme completer were significantly

older and significantly less likely to be reconvicted than the programme dropouts. However, after statistical control of these variables within the logistic regression model, the effect of group on reconviction remained indicating that the differences in risk of reconviction scores and age are not responsible for the observed differences in reconviction outcomes. This chapter, however, was only able to test for differences between the groups' age, number of previous convictions and risk of reconviction scores; some commentators on offender intervention evaluation research have argued that constructs such as motivation and offender need, yet to be measured and controlled for within the body of research, are more likely responsible for the observed effects than the treatment itself (Debidin & Lovbakke, 2005). As such, proponents of such a view propose that "the effect of program completion on reconviction is a product of a nonrandomized design" (Hollin et al., 2008, p.279) rather than the effect of the treatment itself.

The response to such criticism is, perhaps, to consider the notion of a randomised design; a methodology which aims to maximise internal validity and to reduce the influence of unmeasured or unknown factors on observed differences. Within such applied research fields, however, such a design is very difficult to achieve and highly impractical (Colledge, Collier, & Brand, 1999; Farrington, Gottfredson, Sherman, & Welsh, 2002; Hollin, 2008; Pawson & Tilly, 1997; Raynor, 2008). Hollin et al. (2008) have also argued that the effort expended in such research within this field is not necessary given that outcomes of good quality randomised and non-randomised designs rarely differ significantly in their outcomes. As such "tight control within quasi-experimental studies can produce reliable evidence" (Hollin et al., 2008, p.279).

So how can such 'tight controls' be ensured? It is clear from this discussion that an improvement on previous research would be to match the experimental and comparison groups even more closely on key variables. Such a methodology would seek to eliminate the possibility that extraneous variables are responsible for the observed effects. But how are researchers to decide which variables are required to match these groups when it is not known what the influential variable is (or indeed if such a variable exists)? In the present situation, such an exercise would demand a methodology which casts the net wide in the hope of catching the key variable(s). Alternatively, further investigation of the differences between the naturally occurring groups of completers,

non-completers, and non-starters, would determine which variables are routinely associated with group membership and hence which could be responsible for the observed effects. The impact of these variables on observed outcomes, such as reconviction rates, could then be determined by subsequent research. If no such impact were observed, for example, if the effect remained after controlling for these differences, the evidence for a treatment effect would be strengthened.

As previously stated, the aim of this initial study was to add to the research base by evaluating community based accredited programmes within England and Wales using a matched comparison group. It could tentatively be said that the results reported here reflect the findings of other research projects that have utilised statistical control rather than matching of key variables. Cann et al. (2003), Friendship, Blud, Erikson, Travers, and Thornton (2003), Hollin et al. (2004), van Voorhis et al. (2004), Hollin et al. (2008), McGuire et al. (2008) all conducted a 'treatment allocated' analysis within their intervention research and most (with the sole exception of the Friendship et al. study) reported little or no difference between the reconviction rates of those allocated treatment comparative to non-treatment controls. The findings presented here concur with those of previous research in this regard. Further, in relation to the second part of this study, Cann et al. (2003), van Voorhis et al. (2004), Palmer et al. (2007), Hollin et al. (2008), and McGuire et al. (2008) employed a 'treatment received design' which compared outcomes of naturally occurring groups (completers, non-completers) after controlling for population factors. These studies report that completers of interventions tend to have more favourable outcomes than those who failed to complete programmes and/or a comparison group. Notwithstanding the lack of power in the present study, the patterns observed within the odds ratios and effect sizes do reflect these previous findings. The next chapter will attempt a comparison of completer, non-completer and non-starters groups to determine those factors that systematically differ between groups.

Chapter Three

Attrition: A comparison of completers, non-completers, and non-starters

Introduction

As outlined in the concluding paragraph of Chapter two, the present chapter aims to investigate differences between the naturally occurring groups of completers, non-completers and non-starters of community based offending behaviour programmes. The general pattern within the literature is that completers of programmes tend to have more favourable reconviction/recidivism outcomes than comparison groups or programme dropouts (Cann et al., 2003; Friendship et al., 2004; Hollin et al., 2004, 2008; McGuire et al., 2008; van Voorhis et al., 2004). There is also the suggestion within the literature that the outcomes of programme dropouts compare unfavourably to those of comparison groups where statistical control of group differences is employed (Hollin et al., 2008; McGuire et al., 2008; van Voorhis et al., 2004), where the groups are matched on key variables (Hatcher et al., 2008; see also Chapter two) and also within meta-analyses (McMurran & Theodosi, 2007).

The merits of the different evaluation research designs have already been outlined and discussed within chapter two. During this discussion the present author advocated the study of the naturally occurring groups of programme completers, non-completers and non-starters alongside appropriate comparison groups to evaluate treatment impact. However, matching on, or statistically controlling for, variables that differ between the groups is required within such quasi-experimental evaluation research to ensure that any differences between the outcomes of the groups can be attributed as far as possible to the intervention. If such differences are ignored intervention researchers cannot confidently claim that any observed effects are due to the intervention. Even in randomised control trials, attrition is inevitable and it is possible that the natural split of the experimental group into programme completers and dropouts is influenced by variables which also relate to reconviction likelihood. As such, the challenge for research is to first establish by which variables the completer and dropout groups differ and, second, to determine whether it is these variables that are responsible for the effects observed or whether after controlling (or matching) for these variables, a ‘completion’ effect still remains.

More knowledge of the differences between completers and dropouts is also required for operational reasons. In identifying systematic differences between programme completers and dropouts, practitioners would be able to draw on this information to target resources efficiently. Some attempts at this work are evident in the literature: Nunes and Cortoni (2006b) attempted the development of a tool, the Dropout Risk Screen, to aid practitioners target individuals for further assessment and pre-treatment motivational work. Likewise Casey et al. (2007) have developed a self-report questionnaire which, they claim, measures 'treatment readiness' (Ward et al. 2004). This tool, the Corrections Victoria Treatment Readiness Questionnaire (CVTRQ), was able to predict treatment engagement and whilst, to date, there is no research applying this tool to programme attrition[12], there is potential for evaluating differences between completers and dropout groups on this measure.

Nunes and Cortoni (2006b), however, warn against the use of such information to 'cherry pick' for intervention those offenders who are deemed by such tools as most likely to complete. They argue that "a clear *misuse* of the DRS…would be to use it as an indicator of "treatability"" and that "it would be incorrect and inappropriate to refuse an offender entry into a program based on his score on the DRS." (p. iii). Instead they argue that those highlighted by the tool at higher risk of programme dropout should be "targeted with pre-treatment efforts to increase their motivation and general readiness for treatment" (p. iii). Beyko & Wong (2005) echo this view: "results of attrition research should not be used to develop an "attrition profile" to exclude offenders from treatment. Predictors of attrition should be seen as markers for program improvement, rather than shortcomings of the offender." (p. 375).

The research of Nunes and Cortoni (2006b) and Casey et al., (2007), whilst advancing our knowledge of the link between individual factors, attrition and engagement, can be criticised for focussing solely on the individual as the agent of attrition. The factors considered by Nunes and Cortoni (2006b) pertain only to individual variables (risk, age, prosocial attitudes, motivation for intervention and so on) and the questionnaire of Casey et al. (2007) comprises self report items relating attitude and

[12] The author of this thesis is currently (March 2009) engaged in such research. It is being undertaken in conjunction with a local probation area. All offenders referred to a programme and attending a pre-programme interview within the Probation Area are completing this measure. The scores of completers, non-completers, and non-starters of programmes on the CVTRQ will be assessed.

motivation, emotional reactions, offending beliefs and efficacy factors. Despite treatment readiness theory, upon which the CVTRQ was developed, recognising the impact of the "therapeutic situation" (Ward et al., 2005, p. 650) in treatment engagement, there is no consideration within these researches of the external factors that may contribute to attrition or treatment disengagement. Casey et al. do acknowledge this limitation but do not propose how the information gleaned from the CVTRQ could be combined with information relating to the treatment context to create an overarching measure of treatment readiness. Such lack of integration of factors is perhaps due to a general lack of research on the impact of specific responsivity factors on outcome measures such as reconviction, recidivism and programme attrition. Andrews, Bonta and Wormith (2006) in defining specific responsivity as the "matching of service with personality, motivation, and ability and with demographics such as age, gender, and ethnicity" (p. 7) state that "Adherence with specific responsivity and professional discretion has yet to be explored meta-analytically" (p. 12). As such, it is perhaps not surprising that a treatment readiness tool encompassing individual and process or service factors is yet to be developed.

So what do we already know about the differences between completers, non-completers and non-starters of programmes? Chapter one outlined this literature in detail and hence a full review here should not be required. However, in summary, research to date has found relatively robust findings indicating that programme dropouts differ from programme completers by way of risk of reconviction scores (Craissati & Beech, 2001; Hollin et al., 2008; McGuire et al., 2008; Palmer, et al., 2007; Turner, 2006; Wormith & Olver, 2002), criminal history (Babcock & Steiner, 1999; Hollin et al., 2008; McGuire et al., 2008; Palmer, et al., 2007; Zanis et al., 2003), and age (Hazeltine et al, 2002; Hollin et al., 2008; McGuire et al., 2008; Palmer, et al., 2007; Mosher & Phillips, 2001; Robinson, 1995; Turner, 2006; van Voorhis et al,. 2004; Zanis et al., 2003). Further research has provided tentative evidence of differences between the groups in relation to educational level (Babcock & Steiner, 1999; van Voorhis et al. 2004; Wormith & Olver, 2002), marital status (Craissati & Beech, 2001; Miner & Dwyer, 1995; Moore, Bergman, & Knox, 1999; Shaw, Herkov, & Greer, 1995), social class (Miner & Dwyer, 1995), communication skills (Roberts, 2004), literacy skills (Roberts, 2004), and external motivation (Roberts, 2004).

This chapter will attempt to build on the above findings in determining differences between the completer, non-completer and non-starter groups in relation to variables that are routinely available to probation staff. Initially, the research will first assess whether the more robust findings relating to age, number of previous convictions and risk of reconviction are found within this dataset. The analysis will then consider how the probation areas use the information available to them in their selection of participants for programmes.

Community based general offending behaviour programmes have, and they are required for accreditation purposes, evidence based targeting criteria which outline the type of offender most likely to benefit from such intervention. To determine 'eligibility', the programme manuals outline that the offender should have a medium to high risk of reconviction, defined as an OGRS2 risk score of between 31 and 74 (National Probation Directorate, 2001) Previous research, however, has reported that this targeting criterion is not always adhered to by Probation Areas (Palmer et al., 2007, 2008) and that completion levels vary by appropriateness class (too low, appropriate, too high). Such a finding is, perhaps, not surprising given that it is known that completion is associated with risk of reconviction scores. However, Palmer et al. (2008) report large differences; those offenders with OGRS2 score below the lower limit of the criterion were most likely to complete (50%) followed by those who met the criteria (35.53%). Those whose OGRS2 scores were above the recommended limits had the lowest completion rates (23.70%). Given these indications in the literature, the appropriateness of allocation to a general offending behaviour programme will be assessed with respect to programme completion within this chapter.

Once eligibility for a programme is determined, the offender is then assessed for 'suitability'; being suitable for a programme is defined as the offender having cognitive deficits of the type targeted by the intervention (National Probation Directorate, 2001). A 'suitability matrix' is completed by each offender's probation officer and scores above a certain threshold indicate a level of cognitive deficit which would benefit from intervention and hence suitability for the programme. There is very little information within the literature relating to the impact of cognitive deficits on programme completion and the research that is available is inconclusive. Westmarland et al. (nd) reported that programme non-completers had more problems relating to reasoning and

thinking skills than programme completers. However, programme non-starters had fewer needs than non-completers within this area; their profiles matched more closely with the completers than with the non-completers. This chapter will therefore examine whether there are any associations between suitability for the intervention and subsequent dropout, either pre- or during programme.

Another construct discussed within the literature in relation to the potential differences between completers and dropouts of programmes is that of motivation. Despite being an ambiguous concept (Drieschner et al., 2004) and there being a distinct lack of supporting evidence (McGuire et al., 2008), it has been argued that "completion rates are strongly linked to motivation" (Debidin & Lovbakke, 2005). In staking such a claim, however, it is not clear which behavioural outcome these authors consider the object of this motivation to be (Drieschner et al., 2004); is it motivation to stop offending, motivation to attend treatment, or indeed motivation relating to some other behaviour that is strongly linked to completion rates? As McGuire et al. (2008) point out, Debidin and Lovbakke also do not provide an explanation as to how their theory explains the large increases in programme completion rates over recent years. McGuire et al. cites evidence of an increase from 37% of allocated offenders within the community services of England and Wales completing programmes in 2001-2002 to 68% in 2004-5. As such, McGuire et al concludes that: "It is implausible that there could have been a widespread, underlying, improvement in motivational levels of offenders throughout England and Wales" (McGuire et al., 2008, p. 35).

Furthermore, Debidin and Lovbakke argue that findings of quasi-experimental research imply that "the programme simply served to sort those who would do well anyway from those who would not, regardless of the treatment" (p. 48). As such, they place the responsibility for programme completion firmly with the individual and their internal motivational state. In doing so, Debidin and Lovbakke rule out the potential influence of process or organisational factors on programme completion or attrition. This seems to go against the arguments of other commentators within this field. Within the same volume as Debidin and Lovbakke's claims, Chitty (2005) argues that low completion rates reported within the evaluations of offending behaviour programmes within England and Wales (for example, Hollin et al., 2004) are likely due to the failure of probation areas to implement the programmes appropriately. Likewise, following

their qualitative evaluation of the prison-based ETS and R&R programmes, Clarke, Simmonds, and Wydall (2004) claimed that motivation was a factor in programme outcomes but that "motivation for programme participation and individual change are influenced by a complex combination of the individual's characteristics [and] the institutional context in which programmes are delivered" (p. 4). Debidin and Lovbakke's claims would also seem to disagree with the emerging theory of treatment readiness (Ward, et al., 2004) which also acknowledges the contribution of individual and process variables to treatment attendance and engagement. It would seem, therefore, that the weight that Debidin and Lovbakke place on the role of 'motivation' in programme completion and subsequent reconviction outcomes might be misguided. This research will, however, assess the association between programme completion motivation and attrition by testing for differences between programme completers, non-completers and non-starters.

In 2004, the Home Office published an evaluation of the literacy demands of general offending behaviour programmes (Davies, Lewis, Byatt, Purvis, & Cole, 2004). This research assessed the levels of reading and writing required by the three general offending behaviour programmes (Think First, ETS and R&R) and compared these with the skill levels of offenders referred to the programmes. This research concluded that "for many offenders, the literacy demands of the three programmes exceeded their literacy skills" (p. 1). Fifty-seven percent of the offender sample presented with skill levels below that of a competent eleven year old; however, all three programmes required reading skills to be at this level or above. The researchers argued that this is situation is further compounded by the lack of information provided to programme tutors regarding the literacy skills of the offenders on their programmes and, when such needs are identified, the inability of the tutors to be responsive to those needs.

Anecdotal evidence from offenders within the Davies et al. (2004) paper indicates that offenders with literacy needs have problems understanding the vocabulary, language structure, pace, and concepts that comprise the programmes to with they are allocated. Given these findings, it could be hypothesised that those offenders with low literacy levels would be more likely to drop out of a general offending behaviour programme. Indeed, research conducted within West Yorkshire Probation Area (Briggs, Gray, & Stephens, 2004) concluded that at each stage of the ETS

programme (pre-programme psychometrics, programme start, programme completion, post-programme psychometrics) there was greater attrition amongst those with literacy problems than those with no such problems. As such, this research will evaluate whether there is any association between literacy problems, school leaving age, and attrition amongst this national sample of programme referrals.

The final set of variables to be assessed within this chapter is that relating to the needs of the offenders within the sample. The need principle (Andrews & Bonta, 2006) dictates that interventions with offenders which aim to reduce recidivism should assume as their treatment targets those dynamic offender needs associated with an increased likelihood of recidivism. Both the ETS and the Think First programme have been designed with the criminogenic needs of offenders in mind. As such, they focus on the provision of social problem solving skills, self control and self-management skills, the challenging and modification of antisocial values, beliefs and attitudes, and social interaction training (Clark, 2000; Home Office, n.d.).

In recent years within the UK, the OASys risk and need assessment has become the joint prison and probation national tool. However, prior to the full scale implementation of this tool, and at the time at which the data for this research were collected, a scaled down version of OASys was in use in probation areas. The Evaluation and Monitoring Form, or 'mini-OAsys' tool, collected information relating to criminal history and eleven need areas: attitudes towards offending, accommodation, relationships, education and training, employment , finance, lifestyle and associates, alcohol misuse, drugs misuse, emotional factors, and interpersonal behaviour. Debidin and Lovbakke (2005) have argued that "the offenders who completed the programme may have been those less likely to reoffend in any event because they had fewer offence-related needs" (p. 43) and indeed this may be the case. Whilst it has been established through research that the above factors are associated with offending behaviour, there is a lack of evidence as to the contribution of offender need to attrition from programmes. This research therefore aims to test this hypothesis.

As such, this chapter aims to build on the findings of chapter two by testing for differences between the programme completers, non-completers, non-starters on a range of variables. All of these variables relate to the individual offender but this chapter will also assess how these variables have been used by probation areas in

determining eligibility and suitability for attendance on either the ETS or Think First general offending behaviour programme. The findings will be discussed with reference to the implications for programme evaluation and for service delivery.

Hypotheses:

1. Programme completers will be significantly older, have lower risk of reconviction scores and fewer previous convictions that the programme dropout groups.
2. There will be a significant association between group membership (completer, non-completer, non-starter) and OGRS2 appropriateness. The programme dropout groups will contain more inappropriately referred offenders.
3. There will be a significant association between group membership (completer, non-completer, non-starter) and categorical suitability. The programme dropout groups will contain more unsuitable offenders.
4. Programme completers will have significantly lower suitability scores than the programme dropout groups.
5. Programme completers will have significantly higher motivation scores than the programme dropout groups.
6. There will be a significant association between group membership (completer, non-completer, non-starter) and the presence of literacy problems.
7. There will be a significant association between group membership (completer, non-completer, non-starter) and school leaving age.
8. Programme completers will have significantly fewer needs (overall and criminogenic) than the programme dropout groups.

The analyses will also test for associations between group membership (completer, non-completer, non-starter) and the variables of employment status and accommodation status. No hypotheses are made in relation to these variables due to a lack of literature in respect of these relationships.

Method

Procedure and design

This second study used the same national dataset relating to general offending behaviour programmes utilised within chapter two. This second study, however, aimed to determine whether there were systematic differences between completer, non-completer and non-starter groups and hence whether prediction of programme completion/dropout by routinely collected probation data was possible. As such, the comparison group, which had not taken any part in an offending behaviour programme, was not needed for this study. The sample therefore comprised the 173 experimental group participants. This group, during the process of their time of probation, naturally split into the following group: programme completers, non-completers and non-starters. These three groups were compared to determine differences between the groups on key pieces of data routinely collected by Probation Areas.

Participants

Within the experimental group of 173, there were 66 programme completers, 38 non-completers and 69 non-starters. The average age of this sample was 26.92 (*SD* = 7.27) with a range from 18 to 55 years. The mean OGRS2 score was 62.04 (*SD* = 21.63) with a range scores from 8 to 98 and the mean number of previous convictions was 8.66 (*SD* = 8.53) with a range from 0 to 55.

Measures

The data for these analyses were all collected from files held by the Probation Areas. The type of data collected falls into three types: demographic, offender needs and offender programme suitability (or level of cognitive deficit).

Demographic information. The following information (as per chapter two) was collected from probation files or the Offenders Index: the offenders' age, their number of previous convictions and their OGRS2 scores (risk of reconviction). Information was also collected related to the offenders' school leaving age (under 16, aged 16, over 16),

their employment (employed, unemployed) and accommodation (fixed abode, no fixed above) status and whether they had any literacy problems (yes, no). In addition, it was possible to retrieve information relating to the offenders' motivation to complete the intervention to which they were sentenced. This was assessed by the offenders' Probation Officer who was required to answer the question: 'What is the offender's motivation to complete the programme?' Responses were rated on a Likert scale from one (no motivation) to five (high motivation).

Needs scores. Prior to the introduction of the risk and need tool, OASys, Probation Areas were required to complete a paper based scaled down version of OASys, termed 'mini-OASys'. This instrument required Probation staff to assess each offender by completing a series of questions relating to eleven needs:

- attitudes towards offending
- accommodation
- relationships
- education and training
- employment
- finance
- lifestyle and associates
- alcohol misuse
- drugs misuse
- emotional factors
- interpersonal behaviour.

The questions relating to each of these areas aimed to determine whether the offender required any intervention relating to that area. Each question required a response of 0 (no problem), 1 (some problem) or 2 (significant problem). In addition, the assessment tool required Probation staff to indicate whether any of the eleven needs were linked to the probability of the offender re-offending: Is the offender's [insert need] linked to the likelihood of reoffending? This was also scored on three point scale: 0 (unlikely to be linked to reoffending), 1 (possibly linked to reoffending), or 2 (certainly linked to reoffending).

For the purposes of this research, three variables were calculated from the mini-OASys instrument: a 'level of overall' need score, a 'level of criminogenic need' score, and a 'level of non-criminogenic need' score. Initially, and because the number of questions differed between need areas, a standardised score was calculated for each individual need area. To do this, the scores obtained within each need area (omitting the question relating the need to re-offending) were summed. This score was then divided by the number of items within each category to derive a score ranging from 0 to 2, with a score of 0 indicating no need and a score of 2 signifying a large need within this area.

For each need, the score obtained was then classified as either a 'criminogenic' or a 'non-criminogenic' level of need. The need score was deemed to be 'criminogenic' if the answer to the 'likely to be linked to offending' question for that category was 1 (possibly linked to reoffending) or 2 (certainly linked to reoffending). Answers of 0 (unlikely to be linked to reoffending) meant that this need was classified as a 'non-criminogenic' need.

The criminogenic need scores were summed to determine the 'level of criminogenic need' score and the non-criminogenic need scores were summed to determine the 'level of non-criminogenic need'. Finally a 'level of overall need' score was calculated by summing the 'level of criminogenic need' and the 'level of non-criminogenic need' score. The minimum possible score for all of these variables, with eleven need areas scoring from 0 to 2, was hence 0. The maximum possible score was 22.

The only area of need for which the process deviated from the above description was the 'drug use' category. The questions within this category required a different method of need score calculation. The first question within this need area asked whether the offender had ever used drugs. If this question yielded a 'no' response there were no more questions to be answered and a score of 0 was given for this category. If a 'yes' response was given, however, the respondent was then required to indicate, for a list of different drugs, whether they were a non-user, an occasional user, or a heavy user for each of eleven drug types (cocaine/crack, ecstasy, hallucinogenics, opiates, amphetamines, barbiturates, cannabis, benzodiazepines, steroids, solvents, other). If the offender was designated a heavy user in relation to any drug, an overall score of 2

was given for the overall drugs need score. If the offender was not a heavy user of any drug but the initial (has the offender ever used drugs?) question was answered with a yes, then an overall score of 1 was given for the drugs need variable. The above procedure was then followed to determine whether this need was 'criminogenic' or 'non-criminogenic' and the score was added to the level of needs scores in the usual way.

Suitability scores. Another piece of information routinely collected by Probation Areas relates to the suitability of the offender for a general offending behaviour programme. As part of the referral process, a 'suitability matrix' is completed by probation staff. This asks the probation staff whether the offender has problems in the following seven cognitive areas:

- recognising problems
- solving problems
- awareness of consequences
- achieving goals
- understanding other/lacks empathy
- impulsive behaviour
- rigid thinking.

Each cognitive area is scored on a scale of 0 (no problem), 1 (some problem) to 2 (significant problem). The items are summed to produce an overall cognitive deficit score ranging from 0 to 14. As the general offending behaviour programmes aim to increase the cognitive abilities of the attendees, only those with cognitive deficits are deemed to be suitable for the programme. Offenders scoring seven or above[13] are classed as suitable for the programme (National Probation Directorate, 2001).

For the purposes of this research, this cognitive information was collected from Probation Areas and investigated to see if attrition is predictable either from the continuous suitability score or from the dichotomous suitable/not suitable data (those scoring below seven are deemed 'not suitable' and those scoring seven and above,

[13] This is a National Probation Directorate suitability criterion. The cut off of seven indicates that the offender has some problem in relation to all seven cognitive domains OR a significant problem in at least one area (plus some problem in a number of other cognitive domains).

'suitable). For the continuous suitability variable higher scores represented increased levels of cognitive deficits.

Analysis strategy

The first part of this chapter aims to determine whether there were any systematic differences between the groups (completers, non-completers, non-starters) on a range of factors. As such the comparison group utilised in chapter two was dropped from these initial analyses. Initially, univariate analyses were undertaken to assess any significant differences between the three groups. For those variables for which differences were detected, their ability to predict group membership was assessed using multinomial logistic regression models.

The second part of this study utilised the reconviction data relating to the four groups (comparison, completer, non-completer, non-starter). Where the predictive utility of a factor had been determined, it was entered as a control variable in a logistic regression model assessing the utility of group membership to predict reconviction outcome. The 'time at risk' variable was also entered into this model to control for differences in follow up time between the groups (as was done in chapter two). Hence, the 'time at risk' variable was entered at block one, the variable(s) which discriminated the groups in part one of the study was entered at block two, and the group variable (comparison, completer, non-completer, non-starter) was entered at block three. The outcome variable was the dichotomous variable of reconviction (yes/no). If the discriminating variable was responsible for the differences in observed reconviction outcomes, then the addition of the group variable to the model would no longer return a significant result. Conversely, if the group variable still significantly added to the model, it could be claimed that the effect of 'group' still remained even after controlling for observed differences between the groups.

Where possible (and as in chapter two), effect sizes were calculated to determine the magnitude of the relationships between variables.

Results

Demographic analysis:

Age, number of previous convictions and OGRS2: The demographic data were subject to analysis to determine whether the groups differed in respect of these variables. The three groups' mean ages, number of previous convictions and OGRS2 scores are displayed in table 3.1[14].

Table 3.1. Descriptive Statistics by Treatment Groups

	Completers (*N* = 66)	Non-completers (N = 38)	Non-starters (N = 69)
Age	28.38 (7.80)	24.66 (5.64)	26.77 (7.31)
Previous convictions	8.05 (7.13)	7.76 (7.95)	9.75 (9.95)
OGRS2 score	56.38 [a] (22.47 [b])	63.84 (19.22)	66.46 (21.13)

Kolmogorov-Smirnov tests indicated that the parametric assumption of normality was violated for at least one group for the number of previous convictions variable (completers: *K-S Z* = 1.655, *p* = 0.008; non-starters: *K-S Z* = 1.484, *p* = 0.024). Additionally, the homogeneity of variance assumption was violated by the age data (Levene's statistic = 3.210, *p* = 0.043). Kruskall-Wallis tests were therefore used to assess for age and number of previous convictions differences between the groups. Neither of these tests returned significant effects: age, χ^2 (2, *N* = 173) 5.445, *p* = 0.066, Φ_c = 0.18; number of previous convictions, χ^2 (2, *N* = 173) 1.696, *p* = 0.428, Φ_c = 0.10. A one-way ANOVA found significant differences between the groups' OGRS2 scores, F = 3.969, *df* = 2, 172, *p* = 0.021, *r* = 0.21. Post-hoc Scheffe tests revealed significant differences in OGRS2 scores between the Completer and Non-Starter groups (*p* = 0.024) with Completers having significantly lower OGRS2 scores. No other significant pairwise differences were observed.

[14] These are the same statistics given in Table 2.3 on page 55 where differences were assessed between completers, non-completers, non-starters and the comparison group. However, this chapter is testing for differences between the three groups of completers, non-completers, and non-starters only. The data has been reproduced in the above table for ease of reference.

OGRS2 appropriateness: The targeting criteria for the ETS programme state that offenders should have an OGRS2 score between 31 and 74. As such, analyses were undertaken to assess whether these criteria had been adhered to and whether there were any patterns in the data in relation to this. Each offender was classified based on their OGRS2 score into one of two groups: those whose score fell between 31 and 74 ('appropriate') and those who did not ('not appropriate'). Table 3.2 shows the distribution of this classification between the three groups. Almost one in two offenders fell outside of the OGRS2 targeting criteria. Interestingly, relative to both the completers and the non-starters, the non-completers comprised a higher proportion of individuals who fall within the criteria boundaries of 31 to 74. Chi-square analysis found the distribution of appropriate OGRS2 scores between groups to be non-significant, χ^2 (2, *N* = 173) 1.241, *p* = 0.538, Φ_c = 0.08.

Table 3.2. OGRS2 criteria classification by group

	Appropriate (%)	Non appropriate (%)	Total (%)
Completers	38 (38.78)	28 (37.33)	66 (38.15)
Non-completers	24 (24.49)	14 (18.67)	38 (21.97)
Non-starters	36 (36.73)	33 (44.00)	69 (29.88)
Total	98 (100)	75 (100)	173 (100)

Further analysis classified the offenders again into three groups: those who had lower OGRS2 scores than specified by the criteria (30 and below: 'too low'), those who fell within the criteria limits (31 to 74: appropriate), and those who were higher than the criteria (75 and above: 'too high'). Table 3.3 presents the distribution across groups. As can be seen, relative to the non-completer and non-starter groups, the completer group comprises approximately three times as many 'too low' offenders. Additionally, the drop-out groups have a greater proportion of 'too high' offenders comparative to the completers. Additionally, the level of completion within the appropriate and too high groups is much lower than amongst the too low group: only just over a third of those in the appropriate category and one in four of those in the too high category complete the

programme. The difference between these distributions was significant, χ^2 (4, *N* = 173) 10.107, *p* = 0.039, Φ_c = 0.17.

Table 3.3. OGRS2 appropriateness by group

	Too low (%)	Appropriate (%)	Too high (%)	Total (%)
Completers	11 (68.75)	38 (38.78)	17 (28.81)	66 (38.15)
Non-completers	2 (12.50)	24 (24.49)	12 (20.34)	38 (21.97)
Non-starters	3 (18.75)	36 (36.73)	30 (50.85)	69 (29.88)
Total	16 (100)	98 (100)	59 (100)	173 (100)

Motivation to complete the programme: Given the claims in the literature that motivation is a significant factor in programme completion and subsequent reconviction (Debidin & Lovbakke, 2005), it was tested to see whether the groups differed in relation to their motivation to complete the programme. The mean motivation scores for each group are displayed in table 3.4. Kolmogorov-Smirnov tests indicated that the parametric assumption of normality was violated for the motivation variable (completers: *K-S Z* = 1.857, *p* = 0.002; non-starters: *K-S Z* = 1.757, *p* = 0.004) hence non-parametric tests were used to assess for differences amongst the groups. No significant differences were found between the groups: χ^2 (2, *N* = 173) 0.811, *p* = 0.667, Φ_c = 0.07.

Table 3.4. Motivation to complete the programme mean scores by group.

	Motivation (*N* = 173)
Completers	3.21 [a] (0.94 [b])
Non-completers	3.00 (0.96)
Non-starters	3.12 (0.81)

[a] Mean [b] Standard Deviation

School leaving age: Table 3.5 shows the distribution of school leaving ages by group. Data on school leaving age was missing for one completer. This individual was

therefore removed from the analysis. There was no significant association between school leaving age and completion status, χ^2 (2, *N* = 172) = 1.163, *p* = 0.884, Φ_c = 0.08.

Table 3.5. School leaving age by group

	Under 16 (%)	Age 16(%)	Above 16 (%)	Total (%)
Completers	14 (21.5)	41 (63.1)	10 (15.4)	65 (100)
Non-completers	11 (28.9)	23 (60.5)	4 (10.5)	38 (100)
Non-starters	17 (24.6)	41 (59.4)	11 (15.9)	69 (100)
Total	42 (24.4)	105 (61.0)	25 (14.5)	172 (100)

Employment status: Table 3.6 shows the distribution of employment status by group. These data were missing for two completers and one non-starter. These individuals were therefore removed from the analysis. Whilst there were fewer employed individuals within the non-completer and non-starter groups, chi-square analyses revealed no significant association between employment status and completion status, χ^2 (2, *N* = 170) 4.104, *p* = 0.128, Φ_c = 0.15.

Table 3.6. Employment status by group

	Employed (%)	Unemployed (%)	Total (%)
Completers	21 (32.8)	43 (67.2)	64 (100)
Non-completers	9 (23.7)	29 (76.3)	38 (100)
Non-starters	12 (17.6)	56 (82.4)	68 (100)
Total	42 (24.7)	128 (75.3)	170 (100)

Accommodation status: Table 3.7 shows the distribution of accommodation status by group. Data on accommodation status was missing for one completer and one non-starter. These individuals were therefore removed from the analysis. Whilst a greater proportion of non-starters were of no fixed abode, there was no significant

association between accommodation status and completion group, χ^2 (2, *N* = 171) 5.339, *p* = 0.069, Φ_c = 0.18.

Table 3.7. Accommodation status by group

	Fixed abode (%)	No fixed abode (%)	Total (%)
Completers	61 (93.8)	4 (6.2)	65 (100)
Non-completers	37 (97.4)	1 (2.6)	38 (100)
Non-starters	58 (85.3)	10 (14.7)	68 (100)
Total	156 (91.2)	15 (8.8)	171 (100)

Literacy problems: Table 3.8 shows the distribution of literacy problems by group. Data on literacy was missing for one completer and one non-completer. These individuals were therefore removed from the analysis. There was no significant association between literacy and completion status, χ^2 (2, *N* = 171) 0.649, *p* = 0.760, Φ_c = 0.06.

Table 3.8. Literacy problems by group

	Yes (%)	No (%)	Total (%)
Completers	17 (26.2)	48 (73.8)	65 (100)
Non-completers	11 (29.7)	26 (70.3)	37 (100)
Non-starters	16 (23.2)	53 (76.8)	69 (100)
Total	44 (25.7)	127 (74.3)	171 (100)

Needs analysis:

Analyses were conducted to assess for significant differences between the completion groups on the variables of Level of overall needs, Level of criminogenic needs and Level of non-criminogenic needs. Table 3.9 displays the means in relation to these variables broken down by group.

Table 3.9. Need scores by group

	Overall needs	Criminogenic needs	Non-criminogenic needs
Completers	6.24[a] (3.49[b])	5.05 (3.34)	1.19 (1.18)
Non-completers	7.48 (3.28)	5.94 (3.43)	1.54 (1.40)
Non-starters	6.79 (3.39)	5.69 (3.52)	1.10 (1.05)
Total	6.73 (3.39)	5.50 (3.43)	1.23 (1.19)

[a] Mean [b] Standard Deviation

The Kolmogorov-Smirnov one-sample test detected non-normal distributions relating to the Non-starters' Level of non-criminogenic needs score (*K-S Z* = 1.370, *p* = 0.047). This variable was therefore analysed using the Kruskall-Wallis non-parametric test. The mean ranks were Completers = 84.95, Non-completers = 97.43, Non-starters = 83.21. This analysis returned a non-significant effect, χ^2 (2, *N* = 173) 2.206, *p* = 0.332, Φ_c = 0.11.

Two one-way ANOVAs were conducted on the Level of overall needs and Level of criminogenic needs variables. The homogeneity of variance assumption was met for both variables; Level of overall needs, Levene's statistic = 0.187, *p* = 0.830 and Level of criminogenic needs, Levene's statistic = 0.079, *p* = 0.924. Neither analysis indicated significant differences between the groups: Level of overall needs, F = 1.650, *df* = 2, 172, *p* = 0.195, *r* = 0.44; Level of criminogenic needs, F = 0.994, *df* = 2, 172, *p* = 0.372, *r* = 0.11.

The age, OGRS2 and number of previous convictions variables were correlated with the needs variables to assess for any patterns in the data. Beforehand, however, normality tests were run on these variables; the age and number of previous convictions variables were non-normal (*K-S Z* = 1.594, *p* = 0.012 and *K-S Z* = 2.272, *p* < 0.001, respectively). Any correlation including any of these variables (or the Level of non-criminogenic needs variable) was hence computed using Spearman's rho rather than Pearson's r. The correlations are outlined in table 3.10.

Table 3.10. Correlations between age, OGRS2, number of previous convictions and need scores.

	Age	OGRS2	Number of previous convictions
Level of overall need	rho = -0.068	r = 0.164*	rho = 0.102
Level of criminogenic need	rho = -0.020	r = 0.148	rho = 0.142
Level of non-criminogenic need	rho = -0.101	rho = 0.049	rho = -0.027

* $p < 0.05$

Suitability analysis: categorical suitability

On analysis of the suitability variable, 73 (42.2% of the sample) did not meet the suitability criteria for referral to an accredited programme; the remaining 100 (57.8%) did meet the criteria. The mean suitability scores are presented in table 3.11. Comparisons of the two groups reveal significant differences in relation to the groups' OGRS2 scores (t = -2.069, df = 171, p = 0.040) but no differences in relation to age (U = 3624.00, p = 0.936) or number of previous convictions (U = 3075.50, p = 0.077). Suitable offenders had significantly higher OGRS2 scores than non-suitable offenders (64.92 and 58.10 respectively).

Table 3.11. Mean suitability scores by suitability group

	Mean suitability score	OGRS2	Age	Previous Convictions
Not suitable (N = 73)	5.37[a] (2.46[b])	58.10 (21.94)	26.77 (9.65)	7.25 (7.13)
Suitable (N = 100)	9.96 (1.84)	64.92 (21.04)	27.03 (7.54)	9.70 (9.32)
Total (N = 173)	8.02 (3.10)	62.04 (21.60)	26.75 (7.09)	8.60 (8.36)

[a] Mean [b] Standard Deviation

Univariate analysis: Table 3.12 shows the distribution of suitability amongst the groups. This distribution was not significant (χ^2 (2, N = 173) = 3.787, p = 0.151), Φ_c = 0.15.

Table 3.12. Suitability by group (three groups).

	Suitable (%)	Not-suitable (%)	Total (%)
Completers	35 (53.03%)	31 (46.97%)	66 (100%)
Non-completers	19 (50%)	19 (50%)	38 (100%)
Non-starters	46 (66.67%)	23 (33.33%)	69 (100%)
Total	73 (42.20%)	100 (57.80%)	173 (100%)

Suitability analysis: continuous suitability scores

The above analyses were repeated but this time with the continuous suitability score rather than the classification of this variable into suitable or not suitable for programmes. This analysis retains the variance in this variable that is lost in categorising it. Initial tests for normality of this variable found that it was not normally distributed (*K-S Z* = 2.293, $p < 0.001$) hence Spearman's rho correlations were performed between this variable and the age, OGRS2, and previous convictions variables. The only variable with which the suitability score correlated was the OGRS2 score (rho = 0.170, p = 0.025): as the suitability score increased so did the OGRS2 score. The correlations with age and number of previous conviction were non-significant (rho = -0.018, p = 0.812 and rho = 0.135, p = 0.076, respectively).

The mean suitability scores by group are displayed in table 3.13. A Kruskal Wallis test found no significant differences between the groups in relation to their suitability scores (χ^2 (2, N = 173) = 2.012, p = 0.336, Φ_c = 0.11).

Table 3.13. Mean suitability score by group (three group)

	Mean suitability score (SD)	Range of suitability scores
Completers (N = 66)	8.01 (3.28)	0 – 14
Non-completers (N = 38)	7.32 (3.58)	0 – 14
Non-starters (N = 69)	8.42 (2.58)	0 – 14

Summary: From a comparative analysis of the groups (completer, non-completer, non-starter) in relation to a series of demographic variables, criminogenic and non-criminogenic needs variables, and suitability (cognitive deficit) variables, there were only two variables that reliably separated the groups. The first of these was the OGRS2, or risk of reconviction, score and the second was the appropriateness variable (too low, appropriate, too high), the calculation of which is based on the OGRS2 score. The next phase of this study aimed to evaluate whether it is possible to predict, from these variables, the group (completer, non-completer, non-starter) within which offenders would fall.

Prediction of group from variables which significantly vary between groups:

OGRS2 score: A multinomial logistic regression was performed to determine whether it was possible to predict group (completer, non-completer, non-starter) from the OGRS2 scores alone. There was a good model fit on the basis of the entered predictors, χ^2 (148, N = 173) = 166.854, p = 0.138, using a deviance criterion. Comparisons of the log-likelihood ratios for models with and without predictors showed a reliable improvement with the addition of the OGRS2 predictor (χ^2 (2, N = 173) = 7.782, p = 0.020), thus indicating a good model of fit. Correct classification was 53% completers, 0% non-completers and 60.9% non-starters. The overall correct classification was 44.5%, which represents an increase of 11.2% over by-chance accuracy. Analysis of the odds ratios (see table 3.14) indicates that with each unit increase in OGRS2 scores the odds of being in the non-starter group compared to the completer group increase ninefold (p = 0.008). All other comparisons were non-significant.

Table 3.14. Multinomial logistic regression of programme completion as a function of OGRS2

		B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Non-completers vs	Intercept	-1.516	0.627	5.853		
Completers	OGRS2	1.602	0.966	2.754	4.965	0.748 - 32.947
Non-starters vs	Intercept	-1.326	0.547	5.880		
Completers	OGRS2	2.225	0.838	7.059**	9.256	1.793 - 47.791
Non-completers vs	Intercept	-0.191	0.668	0.081		
Non-starters	OGRS2	-0.623	0.982	0.402	0.536	0.078 - 3.675

** $p < 0.01$

OGRS2 appropriateness: The above analysis was repeated but this time the appropriateness classification was entered into the model instead of the OGRS2 score. The Cox and Snell statistic (0.056) indicated that the model did not fit the data well. Comparisons of the log-likelihood ratios for models with and without predictors, however, showed a reliable improvement with the addition of the predictor (χ^2 (4, N = 173) = 9.889, p = 0.042). Correct classification rates were 65.9% for the Completers, 0% for the Non-completers condition, and 34.1% for the Non-starters condition. The overall correct classification was slightly better than the OGRS2 model at 45.7% which represents an increase of 12.4% over by-chance accuracy. Correct classification was 74.2% completers, 0% non-completers, and 43.5% non-starters. Analysis of the odds ratios (see table 3.15) reveal that compared to the too high group, the too low group were less likely to be non-starters than completers (OR = 0.155, p = 0.009). No other pairwise comparisons were significant.

Table 3.15. Multinomial logistic regression of programme completion as a function of OGRS2 appropriateness

		B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Non-completers vs	Intercept	-0.348	0.377	0.853		
Completers	OGRS2 class = 1	-1.356	0.856	2.510	0.258	0.048 – 1.379
	OGRS2 class = 2	-0.111	0.458	0.059	0.895	0.364 – 2.197
	OGRS2 class = 3	0[a]				
Non-starters vs	Intercept	0.568	0.304	3.501		
Completers	OGRS2 class = 1	-1.867	0.719	6.752	0.155**	0.038 – 0.632
	OGRS2 class = 2	-0.622	0.382	2.646	0.537	0.254 – 1.136
	OGRS2 class = 3	0[a]				
Non-completers vs	Intercept	-0.916	0.342	7.196		
Non-starters	OGRS2 class = 1	0.511	0.975	0.275	1.667	0.247 – 11.259
	OGRS2 class = 2	0.511	0.431	1.402	1.667	0.716 – 3.882
	OGRS2 class = 3	0[a]				

[a] This parameter is set to zero because it is redundant

** $p < 0.01$

Does the treatment effect still exist after controlling for the variables which significantly predict completion/dropout? From the above analyses, it has been established that the OGRS2 scores and the OGRS2 appropriateness banding can significantly predict the group (completer, non-completer, non-starter) an offender will fall within. Hence, it may be the case that any observed differences between the groups' reconviction outcomes are due to the underlying differences between the groups' risk of reconviction scores or appropriateness classification. To test this, two separate sequential logistic regression analyses were conducted. These analyses used the full dataset (including the comparison group) as utilised in Chapter 2. As it has already been established (within Chapter 2) that reconviction outcomes can be predicted from group membership, it is not necessary to repeat this analysis. Instead, the analyses to be presented here will determine whether this effect still exists after controlling for these newly established group differences.

In both analyses, to control for already established differences between the groups, the time at risk variable was entered at block one. At block two, the OGRS2 variable was entered in the first analysis and the OGRS2 appropriateness variable was entered in the second to control for the influence of these variables. These variables were not entered together within one analysis due to the assumption of independence of the predictor variables. Finally, the treatment group variable was entered at block three to determine whether the treatment effect observed in Chapter 2 still remained.

OGRS2 score: At block one, as previously observed, the time at risk variable did not produce a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, N = 346) = 19.62, p = 0.012 but the model was significantly better than the constant-only model: χ^2 (1, N = 346) = 18.63, $p < 0.001$. The addition of the OGRS2 score led to a significant improvement in the model: χ^2 (1, N = 346) = 22.557, $p < 0.001$. Finally, the addition of the group variable at block three led to a significant improvement in the model, χ^2 (3, N = 346) = 8.591, p = 0.035 indicating that the observed 'treatment effect' remains after controlling for the OGRS2 variable. The goodness of fit of the model as measured by the Hosmer and Lemeshow Test was good, χ^2 (8, N = 346) = 10.977, p = 0.203, with 69.7% correct classification of cases. This indicates correct prediction from

the model at 19.7% above chance. Table 3.16 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables.

Table 3.16: Logistic regression of reconviction as a function of time at risk, OGRS2, and treatment group (comparison, completer, non-completer, non-starter)

	B	SE	Wald	Exp (B)	95% CI for Exp (B)
Time at risk	0.001	0.001	7.219**	1.001	1.000 – 1.003
OGRS2	2.447	0.579	17.890****	11.554	3.718 – 35.907
Group			8.314*		
Group (1)	-0.559	0.404	1.914	0.572	0.259 – 1.262
Group (2)	-1.033	0.389	7.047**	0.356	0.166 – 0.763
Group (3)	-0.169	0.454	0.139	0.844	0.347 – 2.055
Constant	-1.415	0.556	6.466	0.243	

* $p < 0.05$, ** $p < 0.01$, **** $p < 0.001$

OGRS2 appropriateness: The above analysis was repeated, this time substituting the OGRS2 variable at block two for the OGRS2 appropriateness variable. The block one results are as reported in the logistic regression above. The addition of the OGRS2 appropriateness classification led to a significant improvement in the model: χ^2 (2, N = 346) = 19.543, $p < 0.001$. Finally, the addition of the group variable at block three led to a significant improvement in the model, χ^2 (3, N = 346) = 8.583, p = 0.035 indicating that the observed 'treatment effect' remains after controlling for the OGRS2 variable. The goodness of fit of the model as measured by the Hosmer and Lemeshow Test was good, χ^2 (8, N = 346) = 10.537, p = 0.229, with 69.7% correct classification of cases. Table 3.17 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables.

Table 3.17: Logistic regression of reconviction as a function of time at risk, OGRS2 appropriateness classification, and treatment group (comparison, completer, non-completer, non-starter)

	B	SE	Wald	Exp (B)	95% CI for Exp (B)
Time at risk	0.001	0.001	6.273*	1.001	1.000 – 1.002
OGRS2 class			14.889**		
OGRS2 class (1)	-1.709	0.452	14.283***	0.181	0.075 – 0.439
OGRS2 class (2)	-0.626	0.275	5.184*	0.535	0.312 – 0.917
Group			8.323*		
Group (1)	-0.497	0.402	1.525	0.608	0.276 – 1.339
Group (2)	-1.031	0.388	7.079**	0.357	0.167 – 0.762
Group (3)	-0.133	0.452	0.086	0.876	0.361 – 2.123
Constant	0.648	0.440	2.176	1.913	

* $p < 0.05$, ** $p < 0.01$, **** $p < 0.001$

Discussion

As discussed previously (and replicated in chapter two) findings from good quality quasi-experimental research (Cann et al., 2003; van Voorhis et al., 2004; McMurran & Theodosi, 2007; Palmer et al., 2007; Hatcher et al., 2008; Hollin et al., 2008; and McGuire et al., 2008) have previously suggested two related effects; first, a 'programme completion' effect, characterised by those who complete their intervention having lower reconviction rates than their appropriate comparison groups and, second and more tentatively, a 'programme non-completion' effect, where those who fail to complete programmes are reconvicted at higher rates than their appropriate comparison groups. In the absence of randomisation, however, results such as these should be interpreted with caution: without a randomised design it is not possible to rule out the potential effects of unmeasured variables on the observed outcomes and, arguably, even within randomised designs randomisation is lost once the naturally occurring groups of completers and dropouts are assessed individually. Given this and

the problems associated with conducting randomised trials within an applied setting (Colledge, Collier, & Brand, 1999; Hollin, 2008; Raynor, 2008) another solution is possible. The ability within logistic regression models to control for the influence of measured variables on specified outcomes provides the possibility of ruling out the effect of such variables on the 'completion' or 'non-completion' effect. As the influence of each variable is ruled out in this way, an argument could be constructed that it becomes increasingly likely that the existence of these two effects is a direct result of the intervention rather than an indirect result of pre-existing differences between the groups. Such research also has the potential to contribute to the processes and practices around offending behaviour programmes. The more that is known about programme attrition and the possible reasons for it, the more can be done to attempt to reduce it.

This chapter first aimed to investigate whether there were any significant differences between the three groups (completer, non-completer, non-starter) on any of a set of routinely collected probation-related variables. Where significant differences were established, the predictive utility of these variables was then determined. Finally, those variables found to significantly discriminate between groups were added to the logistic regression model as control variables to assess whether, after the influence of these variables was removed, the effect of group membership on reconviction outcomes (as reported in chapter two) still remained.

Only two of the variables studied here differed significantly between the three groups. From a set of variables including demographic and criminogenic variables (including ratings of motivation to attend a programme), offender need (criminogenic and non-criminogenic) and offender suitability for programmes (or cognitive deficit scores), the two variables which differed significantly were the OGRS2 risk of reconviction score and the associated OGRS2 appropriateness classification (too low, appropriate, too high). Given the findings of previous research (Wormith & Olver, 2002; Hollin et al., 2004, 2008; Palmer et al., 2008), it is not surprising that the completer group had lower risk of reconviction scores than the non-completers and significantly lower scores than the non-starters. However, what might be surprising to some is that nearly half of the sample of offenders sentenced to attend an offending behaviour programme with specified targeting criteria had risk of reconviction scores which placed

them outside of the appropriate risk category for the programme. Despite the Home Office stating that for "OGRS2 scores of 75% or more, only R&R has been demonstrated to have an effect" (Home Office, 2000, p.5), the advice to probation areas at the time of this research was that offenders with OGRS2 scores of 75 and over can attend a general offending behaviour programme but that it should be sequenced with a further intervention, such as another offence focussed offending behaviour programme (National Probation Directorate, 2001). In this respect, the advice recommends that the dosage of the intervention received is increased in line with the risk principle (Andrews & Bonta, 2006). Whilst such advice would fit with the recommendations from research, it is surprising that such a large proportion, over a third, of the sample had risk scores higher than the recommended maximum value.

Such findings, however, are not particular to this research study; Palmer et al. (2008) and Hollin et al. (2004) reported similar patterns within their data; up to a half of those referred to programmes fell outside the appropriateness classification. As found by Palmer et al. (2008), programme completion also varied across the three appropriateness classifications. Similarly, the highest level of completion within the study reported here was found amongst the 'too low' group (68.75%), followed by the 'appropriate' (38.78%), and then the 'too high' group (28.81%). Such a pattern could indicate the existence of a self-selection effect amongst those sentenced to attend a programme; lower risk individuals are seen to be more likely to complete a programme whilst higher risk individuals are seen to be more likely to dropout or fail to complete. Such a pattern alone would account for the observed differences in reconviction outcomes between the groups; lower risk individuals complete programmes and reconvict at lower rates than higher risk individuals who fail to complete programmes and reconvict at higher rates. However, the finding reported here and elsewhere (Hollin et al., 2004, 2008; McGuire et al., 2008; Palmer et al., 2008) that attest to the effect of group (completer, non-completer, non-starter) on reconviction outcomes even after controlling for OGRS2 scores and OGRS2 classification groups, would counter such a position. Likewise, Palmer et al. (2008, p.217) found there to be "a larger effect of treatment completion among the too-high group than would be expected and a larger effect of treatment noncompletion than expected among the too-low group". As such, it

would seem that the data do not return the straightforward results expected if one were to believe the 'would do well anyway' hypothesis (Debidin & Lovbakke, 2005).

Debidin and Lovbakke (2005) have been critical of quasi-experimental research designs for programme evaluation work and have stated that such evaluations have, to date, failed to control for variables which might impact on the observed outcome. In commenting on work completed by Hollin and colleagues (e.g. Hollin et al., 2004) within the probation services of England and Wales, they have stated that:

> Although the research used statistical control of static risk factors, it was unable to discount other plausible explanations of the outcomes arising from dynamic factors, in particular, the effect of selection, in which the offenders who completed the programme may have been those less likely to re-offend in any event because they had fewer offence-related needs or were more motivated to change. (p.42-43)

In response to such criticism, the research reported here tested for differences between the groups in relation to, amongst others, the two factors mentioned here, criminogenic need and motivation to complete a programme. No differences were found between the groups on these constructs. Indeed, no differences were observed in relation to age, number of previous convictions, school leaving age, employment status, accommodation status, literacy, or suitability (as measured by cognitive deficit scores). Given the lack of significance in findings for these constructs, coupled with the enduring ability of treatment group to predict reconviction outcomes even after statistical control of those variables that did systematically differ, it could be argued that the evidence points increasingly towards the existence of a treatment effect. Of course, there are many variables that remain unmeasured, both those relating to the offender, to the process of offending behaviour programme implementation and delivery, and to the contents of the programme itself, but as the influence of more variables on the observed outcomes are eliminated, so the chances of the existence of a treatment effect increase.

The practical implications of this research relate mainly to the selection of appropriate individuals for interventions. On first glance, a completion rate of over two-

thirds within the too low category might appear to justify the decision to allocate these individuals to a programme. However, Palmer et al (2008) report there to be no differences in reconviction outcomes between the programme completers in the 'too low' category and their appropriate comparisons. As such, the programme had little effect on the reconviction outcomes of this group. As the risk principle (Andrews & Dowden, 2006) would dictate, these offenders are of a level of risk of reconviction such that an intensive intervention is not required.

Conversely, it might be assumed that a completion rate of approximately a quarter of the 'too high' category provides evidence against the referral of these individuals to an offending behaviour programme. This finding mirrors those of other researchers (Girard, 1999; Van Voorhis et al., 2004; Wormith & Olver, 2002). Drawing on the data of Palmer et al (2008) again, however, provides evidence of a large effect of programme completion on these individuals: 'too high' programme completers were 44.7% less likely to be reconvicted than the comparison group. Given this information, it would seem to present correctional services with a conundrum: the rate of programme completion amongst high risk offenders is only likely to be one in four, however, those that do complete fare very well in terms of a reduction in reconviction. The question for correctional services to consider is whether the reduction in reconviction amongst the minority of this group outweighs the dropout rate of 75%. Of course, such rates are not necessarily static: further research on the correlates of dropout could inform practices that support larger proportions of high risk offenders through programmes. If the findings reported by Palmer et al. (2008) translate to these offenders, the allocation of such offenders could easily be justified.

A further implication of this research relates to the literacy problems finding. Contrary to the research of Briggs, Gray and Stephens (2004), the rate of literacy problems across completion groups was remarkably constant: approximately one in four offenders, whether completers, non-completers, or non-starters, reported difficulties with reading and writing. Such lack of variation between groups suggests that literacy issues are not an obstacle to completion and hence offenders with such difficulties should continue to be allocated to offending behaviour programmes. Of course, it could be that those offenders with the most severe literacy problems have already been removed from the sample as unsuitable for programmes. Indeed, the percentage of

offenders with literacy problems reported here is low compared to national and international estimates. Within the UK, the Social Exclusion Unit (2002) has reported that 50% of ex-prisoners have poor readings skills and 80% had poor writing skills whilst an evaluation of basic skills amongst probation clients in 2000-2002 reported that a third of the sample had basic skills levels below the level of a competent eleven year old (McMahon, Hall, Hayward, Hudson, & Roberts, 2004). Within the US, estimates of literacy problems have ranged from 50% (Ryan, 1990) to 75% (Herrick, 1991). Despite methodological flaws with this body of research (Rankin, 2005), it would seem that there is a comparatively low rate of literacy problems amongst this sample of offenders referred to a general offending behaviour programme. The reason for this is not clear: the Home Office is clear in its advice that "Literacy and numeracy problems should not be regarded as grounds for exclusion" (National Management Manual, 2000, p. 11) and the Think First Programme Outline (Home Office, n..d) states that:

> Offenders who have learning difficulties, which would prevent a full understanding of the material, should be excluded. This may include offenders with an IQ of 80 or less…Those with literacy difficulties, e.g. poor reading or writing skills should not be excluded. (p. 29)

It may be that in excluding from the programme those with learning difficulties, the rate of literacy problems within the sample reduces also. Alternatively, it could be that the informal assessment of literacy problems within the sample has underestimated the actual occurrence. Further research is required to untangle this issue. What is clear, however, is that within this sample of programme referrals, literacy problems were not found to be associated with programme attrition.

In considering the null effects of a large proportion of the variables tested within this chapter, it would be wise to contemplate the robustness of the measurements utilised. Some of the variables discussed in this chapter represented hard data (e.g. previous convictions, age) or have been subject to extensive evidence based development processes (e.g. OGRS2). However, others are admittedly relatively crude assessments of the construct of interest (e.g. suitability/cognitive deficits, motivation)

and this should be borne in mind when evaluating the results and their implications. For some such variables, alternative measurements might be obvious and available: for example, cognitive deficits could be measured more readily utilising reliable and valid psychometrics as opposed to practitioner views. However, others, such as motivation, are more difficult to assess. As discussed in the introduction to this chapter, the concept of treatment motivation is elusive, ambiguous, and difficult to measure (Driescher, et al., 2004). Nonetheless, the reason for the adoption of these variables for this research was that they are routinely collected by probation services and as such would be accessible to offender managers or programme teams. However, the replication of this research utilising more valid assessments of these constructs would be welcomed.

Another limitation of this research relates to the variety of variables assessed. In evaluating appropriateness and suitability, this chapter has touched very briefly on how the organisational process supporting programmes can impact on attrition. However, this chapter has focused primarily on offender or individual variables. Given the widely documented implementation failure of programmes within the community services of England and Wales and the implied association within this literature between implementation failure and attrition (Hollin et al., 2004; Chitty, 2005; Raynor, 2004), the impact of process variables on dropout should be tested empirically. The second part of this thesis, utilising data collected from one probation area will assess the relationship of attrition with process variables, as well as offender demographic, psychometric, offence related, and need factors.

PART TWO

Investigation of attrition utilising data from one probation area

An Introduction to Part Two

Part one of this thesis investigated attrition from two community based offending behaviour programmes utilising a national sample of offenders. In this investigation, the focus was an examination of differences between the three groups of programme completers, non-completers and non-starters. Where differences were located, the impact of these on the relationship between completion, dropout, and subsequent reconviction was investigated. The only variable found to differ between the groups was the risk of reconviction score as calculated by the OGRS2 tool. When this difference was controlled for, the effect of group (completer, non-completer, non-starter) on reconviction outcomes remained.

Part two of this thesis aims to develop knowledge of the differences between programme completers, non-completers, and non-starters further. The researcher was fortunate and grateful to be granted access to the records of a Probation Area which delivers the Enhanced Thinking Skills (ETS) programme. Over a four month period the researcher collected retrospective information on offenders who had been sentenced to the ETS programme during the years 2005 and 2006. The next three chapters will present an evaluation of the differences between the programme completers, non-completers and non-starters in relation to demographic details, psychometric and attitudinal constructs, offence history, offender need, and process or organisational factors. Some of these factors will already have been assessed in Part One and/or within the literature and hence will be re-evaluated in Part Two to test the robustness of the findings relating to these variables. Other factors, such as the psychometric and process variables' relationships with attrition have received scant coverage within the literature and hence will add to knowledge in this area. The aim of these chapters is therefore to advance knowledge of the differences between programme completers and dropouts. Such information can be useful operationally, but is also interesting on theoretical and empirical levels.

Chapter Four

An investigation of offender demographic and psychometric variables and their relationship with attrition

Introduction

As outlined earlier in this thesis, research assessing the correlates of attrition from offending behaviour programmes has tended to focus on the contribution of individual factors to programme completion or dropout (Craissati & Beech, 2001; Hazeltine et al., 2002; Nunes & Cortoni, 2006a; Nunes & Cortoni, 2006b; Scott, 2004; van Voorhis et al., 2004; Westermarland et al., n.d.; Wormith & Olver, 2002). Indeed the majority of the factors considered by the literature not only relate to the individual but are also static in nature and hence not amenable to change. Such research is useful operationally in helping to identify individuals who may be susceptible to dropout but is also valuable for programme evaluation research; once it is known by which factors the completer and dropout groups differ, it is possible to establish whether there is any influence of these variables on reconviction outcomes independent of programme completion. Where programme completers are seen to reconvict at a lower rate than appropriate comparison groups and in absence of such influence, the evidence stacks up in favour of a completion effect on reconviction outcomes.

As has been discussed and empirically tested within previous chapters, a consensus has emerged within the literature concerning the relationship of programme completion/dropout with offender age and risk of future recidivism. Programme completers tend to be older (Hazeltine et al., 2002; Hollin et al., 2004; Mosher & Phillips, 2001; Robinson, 1995; Turner, 2006; van Voorhis et al., 2004; Zanis et al., 2003), and at a lower risk of further offending behaviour (Craissati & Beech, 2001; Hatcher et al., 2008; Hollin et al., 2004; Turner, 2006; Wormith & Olver, 2002) than programme dropouts. Much less research has assessed the comparative likelihood of the genders failing to complete offending behaviour programmes. Given that the vast majority of convicted offenders are male (National Statistics, 2008), the throughput of women on offending behaviour programmes tends to be small. As such, most of the research within this field has been undertaken with male participants; indeed some programme evaluation researchers have chosen to remove female offenders from the analysis so as not to reduce the validity of their findings for male offenders. The small remaining subset of

females, however, renders it impossible to conduct robust analyses on these data (E. Palmer, personal communication, 2006).

Following an extensive review of the literature, only two papers were found to touch on this issue directly. Pelissier, Camp, and Motivans (2003) found, in their multisite prison research, that female offenders were more likely than male offenders to voluntarily remove themselves from their intervention (although it was not possible for the authors to disentangle programme factors from the influence of gender: there was an indication that the female programmes had "more rigorous requirements than men's programs" (p.13)). Likewise, Hollin et al. (2002b), reporting on a process evaluation of community based offending behaviour programmes within England and Wales, reported that probation staff thought that females were more likely than males to dropout from the programmes to which they were sentenced.

The Home Office contends that offending behaviour programmes delivered within the Probation Service of England and Wales have been designed to be suitable for both male and female offenders (Home Office, 2003a). However, probation staff have reported that they feel that the contents of programmes are not always suitable for female offenders or indeed for minority ethnic groups (Hollin et al., 2002b). David Perry, Head of Interventions within the National Probation Directorate, in 2000, agreed that the needs of female and other minority group offenders may not be adequately addressed by the programmes available at that time:

> There is emerging evidence that women have certain criminogenic needs additional to those of male offenders, and that these need to be addressed. Notable amongst these is the issue of current abusive relationships. Whether there are unique criminogenic needs relating to race which would result in specific programme provision is an unanswered question. (Perry & Johnson, 2000, p. 14)

Such comments have received empirical support: Heilburn, Dematteo, Fretz, Erickson, Yashuhara, and Anumba (2008) reported on an empirical study which compared male and female offenders in respect of their rehabilitation needs. Differences were observed: female offenders were more likely to face financial

difficulty, and their social relationships were more likely to be criminogenic in nature than males. Blanchette and Brown (2006) have also observed that the genders differ in relation to their onset and maintenance of offending behaviour and have described the contribution of social factors, such as financial hardship and violent victimisation, to female offending behaviour. Distinct female needs in the areas of mental health, drug use, family, education and employment have also been highlighted (Bloom, Owen, & Covington, 2003). It would seem therefore that whilst the risk, needs, responsivity model of intervention programming is relevant to female offenders (Dowden & Andrews, 1999), the needs of males and females that require addressing within such interventions appear somewhat distinct in nature.

In addition to the possible suitability issues, the early implementation of accredited programmes in England and Wales saw probation areas struggling with how to process low numbers of female and minority group offenders (Hollin et al., 2002b). The advice from the National Probation Directorate at that time (National Management Manual, 2000) was that lone female and minority ethnic individuals should not be placed within a group: there should be at least two of each minority group within a programme. Such a policy was not always practical, however, and probation areas were often faced with difficult decisions; what should a probation area do if one of the females within the group dropped out? Should the other female remain within the group or should she be withdrawn? Additionally, given that National Standards dictated that all offenders on an order with an associated programme condition should be placed on a group starting within a certain time period (Ministry of Justice, 2007), the probation area often had to choose which of these policies took priority over the other. In reality, there was a mixed picture across probation areas; some took the view that all minority group individuals should be placed on a one to one programme, some of the larger areas attempted to provide programmes specifically for female or ethnic minority groups, and others struggled with the often conflicting advice from the centre (Hollin et al., 2002b). Given these issues, it may be that females and ethnic minority groups are more likely to drop out of offending behaviour programmes than males. The research reported within this chapter will test these hypotheses.

Static factor research might be useful in identifying offenders who are liable to drop out of their programme but its usefulness to practitioners is limited as it does not

provide an indication of how to work with these individuals to increase their chances of completion. Knowing that an older or lower risk offender is more likely to complete their programme might mean that practitioners can cherry pick their referrals so as to increase completion rates (something that commentators within the literature oppose (Beyko & Wong, 2005; Nunes & Cortoni, 2006b)) but such information fails to provide dynamic targets with which practitioners could work to increase an individual's completion likelihood. This chapter will therefore attempt to address this by testing for differences between the completers, non-completers and non-starters on dynamic, as well as static, factors.

The previous chapter provided evidence against the contribution of cognitive deficits to programme dropout: there were no differences observed between completer, non-completer and non-starter groups in relation to the cognitive deficit/suitability scores. The method of measurement utilised was unsophisticated, however, in that it relied on probation officer's assessments of the cognitive deficits of their clients. Prior to the commencement of an offending behaviour programme offenders on a probation order in England and Wales are required to complete an extensive psychometric and attitudinal test battery. The administration of this battery is repeated on programme completion and the scores can hence be used to assess intermediate psychometric and attitudinal change from pre to post programme. The constructs measured include impulsivity, socialisation, locus of control, attitudes towards crime, receptivity to multiculturalism, critical reasoning, identification with a criminal belief system, and social problem solving. Whilst there is a growing body of evidence that shows improvements on such variables from pre to post programme (Blud & Travers, 2001; Blud, Travers, Nugent, & Thornton, 2003; McGuire, 2005; McGuire & Hatcher, 2002; Robinson, Grossman, & Porporino, 1991; Wilson, Attrill, & Nugent, 2003), very little research has assessed for pre-programme differences between completers', non-completers' and non-starters' scores on these psychometrics and attitudinal scales.

The limited research available concludes that programme dropouts tend to present for treatment with more anti-social tendencies than programme completers: hence those who drop out are those who are most in need of intervention. Nooney (2004) compared a national sample of completers and non-completers of the

community based Think First programme and reported that non-completers were more impulsive, had more anti-social attitudes towards crime, had poorer critical reasoning skills, and a more internal locus of control than the programme completers. Nunes and Cortoni (2006b) have also found programme completers to hold more pro-social attitudes than programme dropouts. The present research study aimed to build on these limited findings by comparing the three groups' (completers, non-completer, and non-starters) scores on the psychometric and attitudinal scales available. Given the previous research, it is likely that programme dropouts will have more anti-social attributes and will display stronger anti-social psychometric properties than programme completers.

The aim of the analyses within this chapter, therefore, is to determine whether programme completers, programme non-completers and programme non-starters differ on a number of offender demographic and psychometric constructs. Utilising data collected from one probation area and relating to one offending behaviour programme, the Enhanced Thinking Skills programme (ETS), this chapter will commence with an examination of the relationship between age, risk of reconviction, and attrition to ensure that the patterns observed elsewhere in respect of these variables hold for this sample. The chapter will then move on to assess the associations between gender, ethnic group, attitudes, psychometric variables and attrition. Any differences found between the completers, non-completers, and non-starters will be discussed in relation to the practical and theoretical implications they may present. Chapters five and six will build on the analyses within this and previous chapters by investigating, within the same sample, the relationship between attrition, criminal history and offender need variables (chapter five) and process variables (chapter six).

Hypotheses

Given the research findings discussed above, the hypotheses in relation to this research are as follows:

1. Programme completers will be significantly older than programme non-completers or programme non-starters.

2. Programme completers will be at a significantly lower risk of future reconviction than the programme non-completers and programme non-starters.
3. Female offenders will be significantly more likely to dropout from programmes than male offenders.
4. Ethnic minority group offenders will be significantly more likely to dropout from programmes than white participants.
5. The pre-programme psychometric and attitudinal scores of programme completers will be significantly more pro-social than those of programme dropouts.

Method

Design

This study utilised information held within Probation files to test for differences between, or associations with, the completer, non-completer, and non-starter groups on a number of variables. The independent variable within these analyses was therefore the 'group' which comprised three levels: completer, non-completer, non-starter. There were a number of dependent variables all of which related to offender demographic factors and pre-programme psychometric scores.

Participants

The sample consisted of 293 offenders from one Probation Area sentenced during the period 1st January 2005 to 31st December 2006 to either a community order or a suspended sentence order with the requirement to attend the Enhanced Thinking Skills programme. The Probation Area from which the sample was taken is based within the central region of England and serves a population of approximately one million citizens. The area is comprised of a mixture of urban and rural districts. The programmes unit is based within the area's main city and requires those from the more rural or semi-rural areas to travel to attend the programme in the city.

The average age of those in the study was 25.56 (*SD* = 7.89) with a range of ages from 18 to 55 years. The mean OGRS2 score was 64.42 (*SD* = 21.29) with a range of scores from 10 to 99 and the mean number of previous convictions was 8.27 (*SD* = 7.046) with a range from 0 to 44. Of the sample, 27 (9.2%) were female and 266 (90.8%)

were male. In relation to the offenders' type of sentence, 209 (71.3%) had been sentenced to a community order and 84 (28.7%) a suspended sentence order. Of the sample, 122 (41.6%) failed to commence the ETS programme (non-starters), 45 (15.4%) commenced the programme but failed to complete it (non-completers), and 126 (43%) completed the programme (completers).

Measures

This research utilised data from three centrally developed Probation databases: the Case Recording And Management System (CRAMS), the Offender Assessment System, (OASys), and the Interim Accredited Programmes Software (IAPS). The CRAMS database is used within probation as a case management and monitoring tool. This system was introduced to provide probation officers with the means to "record the details of offenders, sentences, supervision plans and actions taken; to produce some of the reports required on offenders by the courts and others; to produce progress reports to assist in the supervision of offenders; and to produce management information" (National Audit Office, 2001, p. 13). The data retrieved from CRAMS for this analysis were the offenders' date of birth and their sentence date, from which their age at date of sentence was calculated, their gender, and ethnicity.

The OASys database is an offender assessment system which provides "a structured, research-based approach to assessing an offender's likelihood of reconviction, the criminogenic factors associated with offending, and the risk of harm he or she presents" (National Probation Service, 2003b, p. 2). As such, the system holds information relating to the assessment of risk and need for each offender. Within OASys, there is a facility to conduct an OGRS2[15] (Taylor, 1999) risk assessment in addition to a full OASys assessment. For the purpose of the analyses within this chapter, each offender's OGRS2 score was obtained from OASys.

The IAPS database was introduced to Probation Areas as a system to record information relating to the delivery of accredited programmes, such as ETS. As such, the IAPS system records information relating to the offenders' pre and post programme

[15] The OASys tool now calculates OGRS3, an improved version of the OGRS tool (Howard, Francis, Soothill, & Humpreys, 2009). However at the time of data collection, OGRS2 was the risk assessment tool used by Probation Areas and held within OASys and hence this is the information that was collected by the researcher.

psychometric scores, their attendance on the programme to which they are sentenced, their levels of engagement and understanding of the programme, and other such programme related information. The pre-programme psychometrics collected from IAPS for the purposes of this research were those that comprise the General Offending Behaviour Programme Psychometric Battery (National Probation Directorate, 2004). These are:

Eysenck Impulsivity Scale (Adapted from Eysenck & Eysenck, 1978): This measure of impulsivity was amended by the Home Office for use within prison and community settings. As such, the original questions were altered to statements and two items removed. Cronbach's alpha has been calculated on the amended version using a participant sample of 4571 offenders and was found to be good (0.89; National Probation Directorate, 2004). Test-retest reliability was also found to be good (r = 0.86; National Probation Directorate, 2004). High scores on this scale correspond to higher levels of impulsivity.

Gough Socialisation Scale (Gough, 1960): Socialisation has been defined as "the ability or failure to elaborate on an adequate and realistic set of social expectancies and critiques" (National Probation Directorate, 2004, p. 18). Cronbach's alpha has been calculated using the same sample as above and was found to be adequate (0.68; National Probation Directorate, 2004). Test-retest reliability was also found to be good (r = 0.82; National Probation Directorate, 2004). High scores indicate a greater level of socialisation.

Locus of control (Craig, Franklin, & Andrews, 1984): This measures the extent to which the individual believes that events are under their own control. The National Probation Directorate (2004) have stated that offending behaviour programmes are expected to increase a person's internal locus of control as they begin to take more responsibility for their actions. The version used within the battery is an 18-item version. Cronbach's alpha and test-retest reliability have been calculated using the same sample as above and were found to be adequate (0.66 and 0.61 respectively; National Probation Directorate, 2004). Low scores on this variable indicate a more external locus of control and high scores, a more internal locus of control.

Crime PICS II (Frude, Honess, & Maguire, 1994): This provides a measure of attitudes towards crime and criminal behaviour. Although the original scale included four subscales, only three are included within the test battery. These are:

- General Attitudes to Offending: High scores on this variable indicate more anti-social attitudes towards offending.
- Anticipation of Re-offending: High scores on this variable indicate a higher anticipation of re-offending.
- Victim Hurt Denial: High scores on this variable indicate increased levels of denial in relation to the impact of offending on victims.

Cronbach's alpha and test-retest reliability have been calculated on each of these constructs using the same sample as above and were found to be adequate (alpha: 0.87, 0.82, and 0.77 respectively; test-retest: 0.78, 0.76 and 0.60 respectively; National Probation Directorate, 2004).

Quick Discrimination Index (Ponterotto, Burkard, Rieger, Grieger, D'Onofrio, Dubuisson, Heenehan, Millstein, Parisi, Rath, and Sax, 1995): This provides a "general measure of receptivity to multiculturalism" (National Probation Directorate, 2004, p. 19). This measure was incorporated into the battery to provide information relating to perspective taking and rigid thinking. High scores on this scale indicate non-racist and non-sexist attitudes whilst low scores suggest negative attitudes towards ethnic minority individuals and women. The constructs included are:

- Multi-culturalism: this is a measure of cognitive attitudes to racial diversity.
- Racial Intimacy: this is a measure of affective attitudes to more contact with racial diversity.

The scale has been validated on a sample of 220 participants (National Probation Directorate, 2004). Cronbach's alpha was found to be 0.80 and 0.83 for the Multi-culturalism and Racial Intimacy constructs respectively. Test-retest has also been calculated (National Probation Directorate, 2004) and was found to be 0.90 (Multi-culturalism) and 0.82 (Racial Intimacy)

Psychological Inventory of Criminal Thinking Styles (PICTS; Walters, 2002): Two scales of the PICTS have been incorporated into the test battery. These are:

- Current Scale – this scale provide an indication of the individual's identification with a criminal belief system. High scores correspond to a greater identification with this system.
- Cognitive Indolence – high scores on this scale denote low critical reasoning ability and the use of cognitive short cuts in coping with social problems. This is included in the battery as it is thought that programmes can increase the use of critical reasoning skills.

The Current scale has been validated on a sample of American prisoners (National Probation Directorate, 2004). Cronbach's alpha was found to be 0.88. The Cognitive indolence scale has been validated with a UK population of offenders and was found to have a Cronbach's alpha of 0.79 (National Probation Directorate, 2004). Test-retest was found to be 0.73 and 0.86 for the Current scale and Cognitive indolence respectively (National Probation Directorate, 2004).

Social Problem Solving Questionnaire (Freedman, Rosenthal, Donahoe, Schlindt, & McFall, 1978): This requires the reading of a series of life situations and the subsequent ranking of a list of provided possible solutions to the situations in the order of preference from first through to third. The questionnaire comprises three constructs which are calculated through the consideration of the ranked responses. High scores on each of these constructs indicate a greater endorsement of that solution type:

- Assertive problem solving
- Aggressive problem solving
- Passive problem solving.

Validity checks on these constructs were completed with a UK offender population (National Probation Directorate, 2004). Cronbach's alpha was found to be 0.86 (Assertive), 0.87 (Aggressive) and 0.78 (Passive). Test-retest correlations were found to be 0.54 (Assertive), 0.64 (Aggressive) and 0.63 (Passive; National Probation Directorate, 2004).

Procedure

Permission to conduct the research was initially gained from the Probation Area within which the research was to take place. The Assistant Chief Officer with

responsibility for programmes and Programme Manager within the Probation Area were approached by the researcher and the project was discussed at length. Both individuals expressed their feelings that the research would be useful to the Probation Area, as well as to the research base as a whole. Access to the locally held data was therefore granted to the researcher. An application for ethical approval was submitted to the University of Liverpool Committee on Research Ethics (CORE) and this was subsequently approved.

The data were collected by the researcher from four sources of information within the Probation Area. The original dataset was provided to the researcher by the Probation Area's IT Department. This dataset was populated with details of those individuals, sentenced during 2005 or 2006, to any community sentence which contained a requirement to attend the Enhanced Thinking Skills programme within the local Probation Area. The variables within this original dataset permitted the matching of the individual case to data held within the CRAMS, IAPS, and OASys databases using the Case Records Number (CRN), a unique identifier. Each of these databases was accessed manually by the researcher and the required data retrieved. This was subsequently entered into an SPSS spreadsheet to allow data cleaning and statistical analysis.

Data cleaning

The initial SPSS dataset contained records relating to 322 orders with requirements to attend the ETS programme during the defined time period. The dataset was cleaned before analysis and a number of individuals were removed due to incomplete data. It was also necessary to remove duplicates from the dataset. In 24 cases, an individual had been sentenced during the study period to more than one order upon which they were required to attend the ETS programme. To ensure that the data remained independent, it was necessary to remove the duplicates from the dataset. The rules outlined below were followed during this process:

1. Where the orders overlapped, the order upon which the ETS referral was recorded was retained. This tended, in most cases, to be the later order.
2. Where the orders did not overlap and there were no differences in the outcome of the order (for example, the offender had failed to commence the ETS programme on both attempts), the information relating to the earlier order was

retained as this record would be the purest and less likely to be affected by a previous attempt at the programme.

3. Where the orders did not overlap but there were differences in the outcome of the orders, the data relating to the earlier order was again retained. This was for the same reasons as above.

Analysis strategy

This study aimed to build on the findings of the previous chapters in investigating differences between the groups in relation to offender demographic and psychometric variables. Statistical tests were used to test for significant univariate differences between the three groups. Where parametric assumptions were violated, non-parametric equivalent tests were utilised. Where significant differences were found, these variables were entered into a multinomial logistic regression to test for the predictive ability of the statistically significant variables.

Where possible and in line with previous chapters, effect sizes were calculated to determine the magnitude of the relationship between group membership and the variables under assessment.

Results

Offender demographic and psychometric variables:

A series of univariate analyses were conducted to test for significant differences between the three groups (completers, non-completers, non-starters) in relation to offender demographic variables.

Table 4.1: Descriptive Statistics by Group

	Completers (N = 126)	Non-completers (N = 45)	Non-starters (N = 122)
Age	27.56[a] (8.85[b])	22.98 (6.94)	24.45 (6.65)
OGRS2 score	58.44 (21.69)	69.07 (18.76)	68.87 (20.36)

[a] Mean [b] Standard Deviation

Age: The mean ages of the three groups are displayed in table 4.1. As the distribution of this variable was non-normal (completer group, Kolmogorov-Smirnov Z (*K-S Z*) = 1.857, *p* = 0.002, non-completer group, *K-S Z* = 1.706, *p* = 0.006, non-starter group, *K-S Z* = 1.832, *p* = 0.002), the data were analysed using non-parametric tests. The Kruskal-Wallis test revealed a significant difference between the three groups (χ^2 (2, *N* = 293) = 17.849, *p* < 0.001, Φ_c = 0.25). Pairwise comparisons utilising Mann-Whitney U tests revealed significant differences between the completers and non-completers (*U* = 1744, *p* < 0.001) and the completers and non-starters (*U* = 6031, *p* = 0.003). No differences were found between the non-completers and non-starters (*U* = 2239, *p* = 0.067). Completers were hence significantly older than both non-completers and non-starters.

OGRS2: The mean OGRS2 scores per group are displayed in table 4.1 (the specific issue of the number of previous convictions of each group will be examined in Chapter Five). A Kolmogorov-Smirnov test indicated that the data were normally distributed (completer group, Kolmogorov-Smirnov Z (*K-S Z*) = 0.799, *p* = 0.547, non-completer group, *K-S Z* = 0.909, *p* = 0.380, non-starter group, *K-S Z* = 1.166, *p* = 0.132). Levene's test of homogeneity of variance indicated that this assumption was also not violated (Levene's statistic = -.477, p = 0.621). A one-way ANOVA indicated that there were significant differences between the three groups' OGRS2 scores, F = 9.186, *df* = 2, 292, *p* <0.001. Post-hoc Scheffe tests revealed that there were significant differences between the completers and non-completers (*p* = 0.014) and the completers and non-starters (*p* < 0.001). There were no significant differences between the non-completers and non-starters OGRS2 scores (*p* = 0.999).

Gender: Table 4.2 displays the distribution of males and females by group. A chi-square analysis indicated that there was no association between gender and group (χ^2 (2, *N* = 293) = 0.026, *p* = 0.987, Φ_c = 0.01) hence neither gender was more likely to complete or dropout from programmes.

Table 4.2: Gender by Group

	Males (%)	Females (%)	Total (%)
Completers	114 (42.86)	12 (44.44)	126 (43.00)
Non-completers	41 (15.41)	4 (14.81)	45 (15.36)
Non-starters	111 (41.73)	11 (40.74))	122 (41.64)
Total	266 (100.00)	27 (100.00)	293 (100.00)

Ethnicity: Table 4.3 shows the distribution of ethnicity by group. As the sample size did not permit an appropriate number of cases within each cell, a chi-square analysis was not possible on these data. Additionally, it did not make theoretical sense to collapse the categories. Hence the data were observed for any patterns. The vast majority of the sample was of White origin (79.9%), with the second largest category being of Asian origin (10.9%). The ethnicity of the three groups did not vary substantially; each group was comprised of very similar proportions of each ethnic group.

Table 4.3: Ethnicity by Group

	Completers	Non-completers	Non-starters	Total
Asian (%)	18 (56.25)	4 (12.50)	10 (31.25)	32 (100)
Black (%)	5 (38.46)	4 (30.77)	4 (30.77)	13 (100)
Mixed Race (%)	3 (25.00)	2 (16.67)	7 (58.33)	12 (100)
White (%)	98 (41.88)	35 (14.96)	101 (43.16)	234 (100)
Other (%)	1 (100)	0 (0)	0(0)	1 (100)
Not supplied (%)	1 (100)	0(0)	0(0)	1 (100)
Total (%)	126 (43.00)	45 (15.36)	122 (41.64)	293 (100)

Pre-programme psychometrics: Table 4.4 presents the means and standard deviations of the pre-programme psychometric scores by group. The normality assumption was violated for the following variables: Cognitive Indolence (non-

completers: *K-S Z* = 1.725, *p* = 0.005) and Aggressive problem solving (completers: *K-S Z* = 1.716, *p* = 0.006). Additionally, homogeneity of variance was violated for the Assertiveness variable (Levene's statistic = 3.669, *p* = 0.027). For these three variables, hence, non-parametric Kruskal-Wallis tests were used to assess for group differences. A bonferroni adjustment for multiple testing was applied to the interpretation of these data (familywise error rate: 0.05/13 = 0.004).

No significant differences were observed between the groups' Cognitive Indolence (χ^2 (2, *N* = 209) = 0.233, *p* = 0.890, Φ_c = 0.03) or Assertive problem solving scores (χ^2 (2, *N* = 214) = 3.934, *p* = 0.140, Φ_c = 0.14). The comparison of the groups' Aggressive problem solving scores, however, was significant (χ^2 (2, *N* = 212) = 11.144, *p* = 0.004, Φ_c = 0.23). Pairwise comparisons utilising Mann-Whitney U tests revealed significant differences between the completers and non-completers (*U* = 1704, *p* = 0.01) and the completers and non-starters (*U* = 2374.5, *p* = 0.006). No differences were found between the non-completers and non-starters (*U* = 1083, *p* = 0.898). Completers scored significantly lower on the Aggressive problem solving variable than both non-completers and non-starters.

A series of one-way ANOVAs were undertaken on the remaining ten variables. Significant differences were found in relation to the Racial Intimacy scores only, F = 12.209, *df* = 2, 198, *p* <0.001. Post-hoc Scheffe tests revealed that there was a significant difference between the Completers and Non-starters in relation to their Racial Intimacy scores (*p* < 0.001) with the Completers scoring significantly higher than the Non-starters. No differences were observed between the Completers and Non-completers (*p* = 0.297) or the Non-completers and Non-starters (*p* = 0.065). All other comparisons, using a familywise error rate of 0.004, were non-significant (see table 4.6). Indeed such comparisons were also non-significant at p < 0.05, save for Multiculturalism at *p* = 0.017.

Table 4.4: Pre-programme psychometric scores by group

	Completers		Non-completers		Non-starters		Total		*p*
	Mean (SD)	N	Mean (SD)	N	Mean (SD)	M	Mean (SD)	N	
Impulsivity	11.94[a] (5.23[b])	107[c]	12.20 (4.01)	35	12.77 (4.76)	57	12.23 (4.89)	199	0.589
Socialisation	21.20 (5.25)	104	20.57 (4.80)	35	20.43 (5.21)	49	20.88 (5.15)	188	0.637
Locus of control	43.14 (7.17)	115	43.37 (8.45)	40	41.78 (6.31)	55	42.81 (7.21)	210	0.469
General attitudes to offending	34.98 (12.00)	114	35.31 (11.51)	39	36.05 (13.82)	57	35.33 (12.38)	210	0.869
Anticipation of reoffending	11.04 (4.74)	117	11.68 (5.15)	41	11.78 (5.34)	58	11.36 (4.97)	216	0.592
Victim hurt denial	6.26 (3.08)	120	7.05 (3.63)	39	6.40 (3.08)	57	6.44 (3.18)	216	0.402
Multi-culturalism	19.56 (4.44)	114	20.03 (4.60)	35	17.72 (3.87)	53	19.16 (4.40)	202	0.017
Racial intimacy	18.50 (4.34)	112	17.21 (3.78)	34	15.02 (4.25)	53	17.35 (4.46)	199	< 0.001*
Current scale	27.72 (9.03)	115	26.90 (9.84)	39	28.96 (10.80)	52	27.88 (9.63)	206	0.585
Cognitive indolence	18.10 (5.05)	117	19.55 (13.93)	40	17.36 (5.38)	52	18.19 (7.63)	209	0.890
Assertive	53.86 (14.92)	117	48.63 (19.28)	41	48.75 (16.75)	56	51.52 (16.44)	214	0.140
Aggressive	10.02 (8.64)	117	15.41 (12.25)	40	14.83 (11.19)	55	12.28 (10.35)	212	0.004*
Passive	30.74 (9.69)	117	30.35 (10.16)	40	29.46 (8.75)	55	30.33 (9.52)	212	0.716

[a] Mean [b] Standard Deviation [c] Number

Prediction of group from variables which significantly vary between groups:

A multinomial logistic regression was performed to determine whether it was possible to predict group from the age, OGRS2, Aggressive problem solving and Racial Intimacy scores. There was a good model fit on the basis of the entered predictors, χ^2 (372, N = 191) = 335.598, p = 0.912, using a deviance criterion. Comparisons of the log-likelihood ratios for models with and without predictors showed a reliable improvement with the addition of the predictors (χ^2 (8, N = 191) = 33.459, $p < 0.001$). Correct classification rates were 90.9% for the Completers, 3.1% for the Non-completers condition, and 36.7% for the Non-starters condition. The overall correct classification was 62.3%, which represents an increase of 29% over by-chance accuracy. Analysis of the contribution of individual predictors to the model with and without each predictor revealed that the only predictor to significantly predict the outcome was Racial Intimacy (χ^2 (2, N = 191) = 14.851, p = 0.001). All other predictors were non-significant (Age: χ^2 (2, N = 191) = 1.152, p = 0.562; OGRS2: χ^2 (2, N = 191) = 1.881, p = 0.390; Aggressive problem solving: χ^2 (2, N = 191) = 4.447, p = 0.108). Analysis of the odds ratios (see table 4.5) indicates that with each unit increase in Racial Intimacy scores the odds of being in the non-starter group compared to the completer group decrease by 15% ($p < 0.001$) and the odds of being in the non-completers group compared to the non-starter group increase 12% (p = 0.039).

Table 4.5: *Multinomial logistic regression of programme completion as a function of Age, OGRS2, Racial Intimacy and Aggressive.*

		B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Non-	Intercept	-1.247	1.691	0.544		
completers vs.	Age	-0.018	0.032	0.319	0.982	0.923 – 1.045
Completers	OGRS2	0.014	0.011	1.775	1.015	0.993 – 1.036
(ref)	Racial Intimacy	-0.052	0.051	1.037	0.950	0.860 – 1.049
	Aggressive	0.038	0.021	3.182	1.039	0.996 – 1.083
Non-starters vs.	Intercept	1.925	1.485	1.680		
Completers	Age	-0.029	0.029	0.997	0.972	0.918 – 1.028
(ref)	OGRS2	0.006	0.009	0.413	1.006	0.988 – 1.025
	Racial Intimacy	-0.165	0.045	13.320****	0.848	0.776 – 0.926
	Aggressive	0.032	0.019	2.796	1.033	0.994 – 1.073
Non-	Intercept	-3.172	1.899	2.789		
completers vs.	Age	0.011	0.038	0.084	1.011	0.939 – 1.088
Non-starters	OGRS2	0.008	0.012	0.476	1.008	0.985 – 1.033
(ref)	Racial Intimacy	0.114	0.055	4.254*	1.120	1.006 – 1.248
	Aggressive	0.006	0.022	0.066	1.006	0.963 – 1.050

* $p < 0.05$, **** $p < 0.001$

Discussion

The above analyses aimed to investigate offender demographic, attitudinal and psychometric variables to determine whether the three groups of completers, non-completers and non-starters differ in respect of these. Whilst analyses within previous chapters failed to detect statistically significant differences between the groups in relation to their age at index offence[16], the analyses within this chapter identified, in line with the literature (Hazeltine et al., 2002; Hollin et al., 2004; Mosher & Phillips, 2001; Nunes & Cortoni, 2006a; Robinson, 1995; Turner, 2006; van Voorhis et al., 2004; Zanis et al., 2003) that programme completers were significantly older than both non-starters and non-completers of programmes. The groups also differed in respect of their likelihood of future reconviction, as measured by the OGRS2 tool. This finding replicated that reported within chapter three and elsewhere (Craissati & Beech, 2001; Hatcher et al., 2008; Hollin et al., 2004; Turner, 2006; Wormith & Olver, 2002); programme completers had lower risk of reconviction scores than the non-starters and the non-completers. As discussed earlier, such findings could be taken as an indication of a selection effect: those who "would do well anyway" (Debidin & Lovbakke, 2005) are those most likely to complete the programme and then reconvict at a lower rate. However, studies which have statistically controlled for the influence of age and risk of reconviction (Hollin et al., 2008; McGuire et al., 2008; van Voorhis et al., 2004) or have matched the groups on these variables (Hatcher et al., 2008) have reported that the effect of programme completion on reconviction outcomes remains. As such, there appears to be a pattern emerging within the literature that indicates that younger and higher risk individuals are more likely to dropout from community based offending behaviour programmes but that the effect of group (completer, non-completer, non-starter) on reconviction outcomes remains even after the influence of these pre-programme differences between the groups is removed.

[16] Closer inspection of the results reported in the previous chapter reveals that the completers were indeed older than the non-completers and non-starters: the significance level approached the alpha level of 0.05 and the effect size was 0.18. Whilst the age differences between the groups within the present dataset are more pronounced it would seem that the analysis in chapter three was underpowered and that a larger sample size (as is presented here) might have found similar results.

Given the mix of males and females within this dataset, it was possible to assess for an association between gender and attrition of the ETS programme. In contrast to the conclusions of Pelissier, Camp and Motivans (2003) and the anecdotal evidence of Hollin et al. (2002b) which indicated that females were more likely than males to dropout from a programme, these data showed there to be no relationship between gender and programme completion.

In relation to ethnicity, data considerations meant that it was not possible to undertake statistical analyses on these data. Inspection of the frequencies of different ethnic groups across the completion groups indicates that offenders of Asian origin were slightly more likely to complete the programme than other groups. This observation, however, should be considered in context; first, statistical analyses to test for the significance of such associations were not possible, and second, the ethnic composition of the sample is not representative of the national population as the probation area from which these data were collected is situated within an area of England and Wales with a high level of ethnic diversity. Over 20% of the sample comprised non-white individuals; such a proportion is high given the national population statistic of 7.9% (National Statistics, 2006). It should also be borne in mind, however, that the non-white population is overrepresented at all levels of the criminal justice system (Home Affairs Committee, 2007). Therefore it is likely that these figures reflect the convergence of the overrepresentation of minority ethnic groups in the criminal justice system *and* the level of ethnic diversity within the probation area studied.

Given the higher percentage of minority ethnic group offenders within the locality, it is likely that the proportion of non-white offenders within each programme group is larger than would be evidenced in other areas of the country. It is also feasible that the area is able to provide peer and organizational support to such individuals which may not be available within an area of less diversity. Hence the lack of a finding here might speak more to presence of such support within an area of ethnic diversity than to the national pattern of programme dropout amongst non-white sections of the population. As such, a comparative analysis of local probation areas with differing levels of ethnic diversity is required before any firm conclusions can be drawn in relation to this issue.

In an attempt to identify dynamic factors which could be targeted by a pre-programme intervention aimed at reducing attrition, the three groups' pre-programme psychometric and attitudinal scores were compared. The ETS standard test battery returns scores in relation to a host of constructs including impulsivity, socialisation, locus of control, attitudes towards offending, cognitive indolence and so on. Comparisons of the groups' scores revealed that there were significant pre-programme differences in relation to only two of these psychometric constructs: Aggressive problem solving and Racial Intimacy. In relation to Aggressive problem solving, completers were found to be less aggressive in their responses than both the non-completers and non-starters at pre-programme. Such a finding should perhaps be considered in the context the work of Browne, Foreman, and Middleton (1998), Pelissier, Camp and Motivans (2003), and van Voorhis et al. (2004); these authors all independently found that programme dropouts are more likely to have a history of violent offending than programme completers. Howells and Day (2006) have also proposed that violent offenders may have difficulties engaging with interventions due to difficulties in accessing, expressing, disclosing, and reflecting on their affective states. The findings here that programme dropouts tend to be more aggressive would support such a position. Given that high scores on the Aggressive problem solving construct indicate a greater endorsement of aggressive solutions to social problems, it would be useful to assess whether higher scores amongst the dropouts translate into a greater occurrence of violent offending behaviour amongst the dropout sample compared with the completer sample. The following chapter, focused on offence history and offender need, will assess this sample for such a finding. In general, however, the finding that programme dropouts are more aggressive within social situations than programme completers reiterates the OGRS2 finding reported above; programme dropouts have a greater need for intervention than programme completers.

Completers also had significantly higher Racial Intimacy scores than the non-starters, indicating that they hold more positive affective attitudes towards racial diversity than the non-starters (Ponterotto, Potere, & Johanson, 2002). Indeed the logistic regression assessing the predictive ability of all those variables which differed significantly between the groups showed that the Racial Intimacy variable was the only one to reliably predict group membership. The Quick Discrimination Index from which

this construct is taken is “a measure of attitudes underlying potential discriminatory behavior” (Ponterotto, Potere, & Johansen, 2002, p. 193) towards ethnic minorities and the Racial Intimacy construct itself measures “personal, affective comfort with interracial interactions” (p. 194). Despite not citing evidence for such claims, the General Offending Behaviour Programmes Evaluation Manual (National Probation Directorate, 2004) reports that this construct is “relevant to the aims of the General Offender/Cognitive Skills Programmes… [as] the attitudes found to be associated with racism share communality in the programme target areas of rigid thinking and negative attribution to other groups, as well as the location of blame in external events” (p. 19-20). Elsewhere Racial Intimacy scores have found to correlate with completion of course work, uptake of continuing education, attendance at in-service diversity training, and intellectual curiosity measured through travel, book reading, and movie viewing (King, 1997, cited in Ponterotti, Potere, & Johanson, 2002).

From the findings of King (1997, cited in Ponterotti, Potere, & Johanson, 2002) and the proposition of the National Probation Directorate (2004), it could be hypothesised that the Racial Intimacy construct is just one component of a group of constructs that are strongly inter-correlated amongst programme non-starters. These constructs would include aggressiveness, rigid thinking, negative attributions to out-groups (including minority ethnic groups), an external locus of control, a lack of engagement in educational programmes and low levels of intellectual curiosity. Previous research into the pre-programme psychometric differences between programme completers and drop-outs is far from plentiful and most has not reached peer-review publication. However, Nooney (2004) reported on an evaluation of Think First within Probation Areas of England and Wales within which pre-programme completer and non-completer (as opposed to non-starter) data were available. Programme non-completers were found to be more impulsive, hold more criminogenic attitudes to offending, have a more external locus of control, and higher levels of cognitive indolence than the programme completers.

Whilst the conclusions drawn from this evidence can be nothing more than tentative, there is an indication that programme attrition is influenced by an underlying construct, such as an “antisocial personality pattern” (Andrews & Bonta, 2006, p. 67). One of the Big Four risk/need factors which Andrews and Bonta argue are predictive of

criminal behaviour in individuals, the antisocial personality pattern is characterised, in personality terms (and using the Five Factor Model; Digman, 1990), as one that is low in agreeableness and conscientiousness or, in relation to dynamic need, as displaying weak self control, weak anger management, and poor problem solving skills. It is conceivable that the findings reported here could indicate that programme dropouts have a more antisocial personality than programme completers.

Such an argument, however, could be construed as lending supportive to the self-selection argument of commentators such as Debidin and Lovbakke (2005). The personality variables known to be associated with reconviction are also those that separate out programme completers from programme dropouts. As such, some would argue that the search for differences between programme completers and dropouts becomes futile: it will merely serve to unearth variables that contribute to the calculation of risk. In such a position, risk predictors will remain the best predictors of both programme completion and outcomes.

This may be so; however, such a position is not necessarily analogous to saying that the programme has no effect of programme completers. As discussed elsewhere, the self-selection hypothesis would maintain that if programme attrition and reconviction are predictable from the same variables, all the programme therefore serves to do is to separate out those that will go on to reconvict (the dropouts) from those who will not (the completers). However, in situations where the first element of this hypothesis has been found to hold, there is evidence that the second element does not necessarily follow. As reported earlier in this thesis, risk of reconviction scores predict both reconviction outcomes and attrition from programmes. However, programme completion was found to predict reconviction outcomes independently of risk of reconviction scores. Likewise, the evidence from Palmer et al (2008) and Palmer et al (2009) would contradict such a position: high risk offenders who complete programmes return large reductions in reconviction outcomes. If the programme has no impact on recidivistic behaviour, these offenders would be destined to reconvict at higher rates than what is observed. It would seem, therefore, that reconviction and attrition may be influenced by similar underlying factors but that it does not necessarily follow that the intervention has no impact on reconviction rates; interventions can still impact on the subsequent behaviour of those who complete the full dosage. The

implication of this in relation to completion rates is that services should be enabled to identify those individuals with high-end antisocial personalities and provided with the resources to support these individuals through their programme.

The next set of analyses, to be presented in the following chapter, will investigate the offenders' offence related and criminogenic factors to determine whether there are any differences between the groups. This information includes details of their index offence, their criminal history, their criminogenic and non-criminogenic needs, and their calculated risk of reconviction.

Chapter Five

An investigation of offence-related variables, criminogenic and non-criminogenic need and their relationship with attrition

Introduction

As outlined in the previous chapter and building on the assessment of offender demographics, psychometric and attitudinal measures, this chapter aims to investigate the relationship between attrition, offence related variables, and offender need variables. As has been established elsewhere (Craissati & Beech, 2001; Hollin, et al., 2004; Palmer, et al., 2008; Turner, 2006; Wormith & Olver, 2002) and within this thesis, there is an established link between attrition and risk of reconviction scores: those with higher risk scores are more likely to fail to complete the programme to which they have been referred. Given that most actuarial assessment tools utilise variables relating to the individual's previous criminal history, such as the number and type of any previous convictions, it is therefore relevant to investigate which of the variables that contribute to the calculation of risk are related to completion of, or dropout from, programmes. This chapter will therefore examine the items that comprise the OGRS2 risk of reconviction score to determine which of these variables may contribute to the differences in risk scores between the attrition groups.

The OGRS2 risk of reconviction score is determined through a statistical algorithm using the following criminal history variables: index offence, history of breach of probation, parole, license, bail or a community based sentence, number of previous convictions prior to age 18 and since the age of 18, age at first conviction, history of custody, diversity of previous convictions (number of different types of offences), and type of previous convictions (whether there is the presence of violent and/or sex offences in the offence history) (Taylor, 1999). Using a large dataset relating to a previous cohort of offenders, the OGRS2 tool was developed by assessing which variables correlated with reconvictions within a two year period (Taylor, 1999). The above listed variables were found to be the most robust predictors and hence were weighted appropriately within the algorithm that is now used extensively within the National Offender Management Service[17]. Whilst the relationship of these variables with

[17] As mentioned previously within this thesis, the National Offender Management Service have recently introduced the OGRS3 risk of reconviction tool. However, at the time of data collection the tool within circulation was the OGRS2 tool. The data collection therefore reflects this.

reconviction is therefore known, it is yet to be investigated how or whether the variables that contribute to the OGRS2 calculations are associated with treatment completion within community settings.

As discussed previously within this thesis, the 'would do well anyway' argument (Debidin & Lovbakke, 2005) asserts that those who complete programmes are those that would otherwise have fared well in terms of their reconviction outcomes. As such, the argument disputes that the medium of change in relation to lower reconviction rates amongst completers compared to dropouts and comparison groups is the intervention itself. Instead the 'would do well anyway' argument claims that change is likely due to the offenders' motivational states and that the programme itself plays no role in the process. It would follow therefore that if the programme serves only to separate those who would do well in terms of reconviction from those that would not, those factors that have been shown to be related to reconviction would also predict programme completion. If the 'would do well anyway' viewpoint holds, therefore, it would be expected that all of the OGRS2 variables, having been selected due to their relationship with future reconviction, would also hold relationships with attrition. This chapter will therefore investigate the OGRS2 variables and their relationship with programme completion in an effort to understand more about the correlates of attrition and consequently the efficacy of interventions for offenders.

The second aim of this chapter is to further investigate the relationship of offender needs with attrition from offending behaviour programmes. The need principle (Andrews & Bonta, 2006) dictates that the most effective interventions will be those which address the criminogenic needs of offenders. It is therefore logical that the population of offenders targeted for general offending behaviour programmes will present with a number of needs originating from variety of domains within their lives. How these relate to attrition is, however, as yet unknown. Anecdotal evidence from probation officers' reports of attrition indicate that the more chaotic a lifestyle an offender has, the less likely he/she is to complete a programme (Hollin et al., 2002a; 2002b), but as yet there is no empirical research testing such hypotheses.

Chapter three presented an evaluation of the relationship between offender need (overall, criminogenic and non-criminogenic) and attrition and found no association between these variables and the likelihood of programme dropout. These

data, however, were derived from crude measurements of offender need using a tool, the 'mini-OASys', which was implemented within some Probation Areas as a precursor to the full OASys risk and need tool. The analyses within this chapter will use data collected from the full OASys tool and will repeat the previous analyses and develop them further. As such these analyses will look not only at the relationship between attrition and overall, criminogenic, and non-criminogenic need, but will also investigate whether any particular need domains contribute to attrition from offending behaviour programmes. As with the discussion of offender risk, if the 'would do well anyway' argument is valid, it would be expected that the overall OASys score, which itself is a predictor of future reconviction (Howard, 2006), should also be associated with attrition: the programme dropouts would have higher scores than programme completers. Additionally, it would be expected that those need domains deemed to be criminogenic, and hence by their very nature linked with the offenders' likelihood of future offending behaviour, would also predict attrition from offending behaviour programmes.

Hypotheses

1. There will be a significant association between group (completer, non-completer and non-starter) and previous breach of probation/parole/license/bail/community based sentence: programme dropouts will be more likely to have breached a previous order than programme completers.
2. Programme dropouts will have significantly more previous convictions than programme completers.
3. Programme dropouts will have a significantly younger age at first conviction than programme completers.
4. Programme dropouts will have significantly more custodial sentences than programme completers.
5. Programme dropouts will have significantly more needs than programme completers.
6. Programme dropouts will have significantly more identified criminogenic needs than programme completers.

In addition to testing the above hypotheses, the analysis within this chapter will contain some exploratory investigation into the relationship between attrition, index offence, and other previous offence types. In addition, the level of need within particular domains (e.g., financial or accommodation) between the three groups will be investigated.

Method

Design

This study aimed to determine whether the groups (completer, non-completer, non-starter) differed in respect to their offence-related variables, or their criminogenic and non-criminogenic needs. The independent variable within these analyses, hence, was the group which again comprised three levels: completer, non-completer, non-starter. The dependent variables this time were all offence related or criminogenic and non-criminogenic needs variables.

Participants

As in the previous chapter, the participants were 293 offenders from one Probation Area sentenced between 1st January 2005 and the 31st December 2006 to a community sentence with the addition requirement of attendance at the ETS offending behaviour programme. The participant descriptives were therefore the same as within the previous chapter: the average age of those in the study was 25.56 (*SD* = 7.89) with a range of ages from 18 to 55 years; the mean OGRS2 score was 64.42 (*SD* = 21.29) with a range scores from 10 to 99; and the mean number of previous convictions was 8.27 (*SD* = 7.046) with a range from 0 to 44. Of the sample, 27 (9.2%) were female and 266 (90.8%) were male. In relation to the offenders' type of sentence, 209 (71.3%) had been sentenced to a community order and 84 (28.7%) a suspended sentence order. Of the sample, 122 (41.6%) failed to commence the ETS programme (non-starters), 45 (15.4%) commenced the programme but failed to complete it (non-completers), and 126 (43%) completed the programme (completers).

Measures:

As explained within the previous chapter, data were obtained from a variety of Probation electronic resources: CRAMS, OASys, and IAPS. The CRAMS database provided the offenders' index offence information, whilst the IAPS system allowed classification of the offenders into programme completers, non-completers, and non-starters. The majority of the data for the analyses within this chapter, however, was extracted from the OASys database. The full OASys assessment requires the collection of data relating to the offenders' offending history and their current offence, their social and economic circumstances (for example, their access to accommodation, education, training and employability, financial management and income, lifestyle and associates, relationships, and drug and/or alcohol misuse) and personal factors (such as their thinking and behaviour skills, their attitudes towards offending and towards supervision, and emotional factors such as anxiety or depression). This information is used by Probation staff to predict the likelihood of the offender being reconvicted and to inform the offender's sentence plan. Offence-related and offender need data were collected by the researcher from the offenders' full OASys assessments. OASys assessments are repeated at various points within an offender's sentence to assess dynamic changes in risk and need. The assessment selected for this research was that completed to inform the pre-sentence report relating to the offender's index offence. In the small minority of cases where this assessment was not available, the assessment immediately following the sentence date was instead selected for data collection.

Procedure:

The procedure relating to the collection of the data has been outlined in detail within the previous chapter. In most cases, the dependent variables were taken directly from the CRAMS, IAPS and OASys databases and used in their raw status within the following analyses. However, the OASys data were used by the researcher to calculate individual criminogenic and non-criminogenic need scores for each participant. These were calculated using the eleven need domain scores which, along with the criminal history scores, comprise the overall OASys assessment. Within the OASys assessment and for each of these eleven need domains, there is a question: for example, 'Are accommodation issues linked to the likelihood of future offending behaviour'. This

question is answered by indicating whether the response is 'yes' or 'no'. Using the responses to these questions, it was possible to determine whether that particular need domain was linked *for that particular individual* to their likelihood of reoffending. In such cases, this was classified as a criminogenic need; where it was not linked to the individual's likelihood of reoffending it was classified as a non-criminogenic need. An overall criminogenic need score was obtained by summing the domain scores identified as being criminogenic whilst the overall non-criminogenic need scores was obtained by summing the domain scores identified as non-criminogenic. Finally, the total needs score utilised the overall OASys score (i.e. the sum of the criminogenic and non-criminogenic need scores). Hence for each individual the dataset contained a total needs score, an overall criminogenic need score, an overall non-criminogenic need score and eleven domain scores identified on an individual level as being either criminogenic or non-criminogenic.

In line with the chapter two discussions which considered the appropriateness of parametric analyses with the OGRS2 data, it was necessary to consider whether these analyses were appropriate for use with the OASys data. The total OASys score (from which the criminogenic and non-criminogenic scores are derived) represents a prediction of "the likelihood of reconviction for each offender". As such, one offender's score on this scale does not alter another offender's score and hence the data meet the parametric assumption of independence (Field, 2005). Additionally, the assumption of interval data is also met: the scores represent a scale of how likely a person is to be reconvicted: the higher the score, the higher the likelihood. These scores are derived from the summing of need area scores which have been weighted in accordance with their link with reconviction. As such, it is the intention that scores on this scale represent interval data in that equal distances on the scale represent equal measurements of the risk of reconviction. The final two parametric assumptions, those of normality and homogeneity of variance, are assessed for each sample utilised. These analyses are reported in the results section as appropriate. Where both these assumptions are met, parametric analyses are utilised; where either assumption is not met, non-parametric tests are used instead

Results

Offender offence-related variables:

Index offence: The index offences of each participant were coded into offence categories by the researcher using the Home Office offence category coding scheme. This coding results in ten offence categories: violence against the person, sexual offences, burglary, robbery, theft and handling stolen goods, fraud and forgery, criminal damage, drug offences, other (excluding motoring offences), and motoring offences. The distribution of offence categories across the sample is displayed in table 5.1.

Table 5.1: Distribution of offence categories for index offences across the sample.

Offence category	N	%
Motoring offences	78	26.62
Theft and handling stolen goods	65	22.18
Violence against the person	45	15.36
Other (excluding motoring offences)	44	15.02
Burglary	23	7.85
Drug offences	17	5.80
Fraud and forgery	11	3.75
Criminal damage	9	3.07
Robbery	1	0.34
Sexual offences	0	0
Total	293	100.00

There were no offenders whose index offence was of a sexual nature within the sample. The predominant classifications were motoring offences and offences related to theft and handling stolen goods.

The distribution of offence categories across groups is displayed in table 5.2. Non-completers comprised almost twice the proportion of individuals who had committed violence offences than did the non-starters. The non-starters had a markedly

greater proportion of individuals convicted of criminal damage offences than the other two groups. Completers comprised a greater proportion of individuals who had committed offences within the "Other" category than either the non-starters or non-completers.

Table 5.2: Distribution of offence categories of index offences across groups (%).

Offence category	Completers	Non-completers	Non-starters
Motoring offences	30 (23.81)	11 (24.44)	37 (30.33)
Theft and handling stolen goods	29 (23.02)	8 (17.78)	28 (22.95)
Other (excluding motoring offences)	23 (18.25)	5 (11.11)	16 (13.11)
Violence against the person	22 (17.46)	10 (22.22)	13 (10.66)
Drug offences	8 (6.35)	3 (6.67)	6 (4.92)
Burglary	7 (5.56)	5 (11.11)	11 (9.02)
Fraud and forgery	6 (4.76)	1 (2.22)	4 (3.28)
Criminal damage	1 (0.79)	1 (2.22)	7(5.74)
Sexual offences	0 (0)	0 (0)	0 (0)
Robbery	0 (0)	1 (2.22)	0 (0)
Total	126 (100)	45 (100)	122 (100)

A chi-square analysis of this distribution (without the sexual offences row as no data were present) returned eleven cells (40.7%) with expected counts of less than five, meaning that the use of chi-square on these data was unreliable as more than 20% of cells had expected counts that were too low (Tabachnick & Fidell, 2006). To attempt to resolve this, the offence categories of burglary, robbery, theft and handling stolen goods, and fraud and forgery were collapsed into one category entitled 'acquisitive crime'. [18] This classification, however, still returned four cells (22.2%) with expected

[18] There is a debate as to whether the crimes of fraud and forgery should be classified within the category of acquisitive crime alongside crimes such as burglary and robbery. For example, Dodd (1998) concluded that only 13% of the 209 fraudulent insurance claim perpetrators he examined were in financial difficulty and 57% were earning a regular income; such findings may indicate different motivations for offending and hence different offender types. It was decided, however, to include these crimes within this classification based on the definitions of acquisitive crime used elsewhere in the academic literature (e.g.

counts below five. Hence rather than undertake a 3 (completer, non-completer, non-starter) x 6 (offence codes) chi square, the analysis was conducted on pairwise comparisons of the groups, for instance, completers versus non-completers, completers versus non-starters, and non-starters versus non-completers. As such, three 2 (completion group) x 6 (offence codes) were undertaken. This solution resulted in fewer cells with low expected counts within each analysis: indeed, none of these analyses exceeded 20% of cells with expected counts less than five. Table 5.3 presents the distributions upon which these analyses were undertaken.

Table 5.3: *Distribution of offence categories of index offences (with the collapsed category of acquisitive crime) across groups (%).*

Offence category	Completers	Non-completers	Non-starters
Acquisitive crime	42 (33.33)	15 (33.33)	43 (35.25)
Motoring offences	30 (23.81)	11 (24.44)	37 (30.33)
Other (excluding motoring offences)	23 (18.25)	5 (11.11)	16 (13.11)
Violence against the person	22 (17.46)	10 (22.22)	13 (10.66)
Drug offences	8 (6.35)	3 (6.67)	6 (4.92)
Criminal damage	1 (0.79)	1 (2.22)	7(5.74)
Total	126 (100)	45 (100)	122 (100)

All chi-squares were non-significant: Completers versus Non-completers (χ^2 (5, N = 171) = 2.2024, p = 0.846, Φ_c = 0.11); Completers versus Non-starters (χ^2 (5, N = 248) = 9.037, p = 0.108, Φ_c = 0.19); and Non-completers versus Non-starters (χ^2 (5, N = 167) = 4.763, p = 0.445, Φ_c = 0.17). Thus, there is no evidence of between-group differences in index offence type.

Previous breach of probation/parole/license/bail/community based sentence: Table 5.4 shows the number of offenders within the sample who have previously breached a criminal justice order (probation/parole/license/bail/community based

Gossop, Marsden, Stewart, & Rolfe, 2000; Stewart, Gossop, Marsden, & Rolfe, 2000; Van Der Zanden, Dijkgraaf, Blanken, Van Ree, & Van Den Brink, 2007).

sentence). Analysis of these data shows that almost two-thirds of the sample has previously been breached. Completers have the lowest rate of breach and non-starters have the highest. This association between group and breach was significant (χ^2 (2, N = 293) = 6.542, p = 0.038, Φ_c = 0.15).

Table 5.4: Breach by group

	Yes	No	Total
Completers (%)	69 (54.8)	57 (45.2)	126 (100)
Non-completers (%)	28 (62.2)	17 (37.8)	45 (100)
Non-starters (%)	86 (70.5)	36 (29.5)	122 (100)
Total (%)	183 (62.5)	110 (37.5)	293 (100)

Previous convictions (under/over 18/total): Table 5.5 presents details relating to the mean numbers of previous court appearances at which convicted, below the age of 18 years, from 18 years onwards and in total broken down by group. Kolmogorov-Smirnov tests revealed that these data were non-normal (Convictions under 18: completers: *K-S Z* = 2.520, $p < 0.001$, non-completers: *K-S Z* = 1.512, $p = 0.021$; non-starters: *K-S Z* = 2.104, $p < 0.001$; Convictions 18 and above: completers: *K-S Z* = 2.098, $p < 0.001$, non-completers: *K-S Z* = 1.944, $p = 0.001$; non-starters: *K-S Z* = 2.247, $p < 0.001$; Total convictions: completers: *K-S Z* = 2.023, $p = 0.001$, non-completers: *K-S Z* = 1.450, $p = 0.030$; non-starters: *K-S Z* = 1.605, $p = 0.012$). Three Kruskal-Wallis tests were therefore undertaken to test for significant differences between the three groups. Significant differences were observed in relation to the Convictions under 18 variable, χ^2 (2, N = 293) = 9.293, p = 0.01, Φ_c = 0.18. Pairwise comparisons utilising Mann-Whitney U tests revealed significant differences between the completers and non-starters (U = 6001.5, p = 0.002). No differences were found between the completers and non-completers (U = 2473, p = 0.186) or the non-completers and non-starters (U = 2536.5, p = 0.444). Completers, therefore, had significantly fewer court appearances at which they were convicted under the age of 18 comparable to the non-starters.

Table 5.5: Mean number of court appearances at which convicted by group

	Convictions under 18	Convictions 18 and above	Total convictions
Completers	1.87[a] (2.46[b])	5.64 (5.80)	7.51 (6.41)
Non-completers	2.96 (3.91)	5.09 (6.63)	8.04 (8.40)
Non-starters	3.16 (3.60)	5.98 (6.69)	9.13 (7.10)
Total	2.57 (3.27)	5.70 (6.30)	8.27 (7.05)

[a] Mean [b] Standard Deviation

No significant differences were observed between the three groups in relation to the Convictions 18 and above (χ^2 (2, N = 293) = 0.540, p = 0.763, Φ_c = 0.04) or the Total number of convictions (χ^2 (2, N = 293) = 5.035, p = 0.081, Φ_c = 0.13).

Age at first conviction: Table 5.6 presents the mean ages of first conviction by group. Kolmogorov-Smirnov tests revealed that these data were non-normal (completers: *K-S Z* = 1.917, p = 0.001, non-starters: *K-S Z* = 1.850, p = 0.002). A Kruskal-Wallis test was therefore employed to test for differences between the groups. A significant difference was observed: χ^2 (2, N = 293) = 8.500, p = 0.014, Φ_c = 0.17. Pairwise comparisons utilising Mann-Whitney U tests revealed significant differences between the completers and non-starters (U = 6100, p = 0.005). No differences were found between the completers and non-completers (U = 2528.5, p = 0.280) or the non-completers and non-starters (U = 2364, p = 0.166). Completers were significantly older than the non-starters at their first conviction.

Table 5.6: Mean age of first conviction by group.

	Age at first conviction
Completers	18.27[a] (5.65[b])
Non-completers	16.80 (2.19)
Non-starters	16.51 (3.15)
Total	17.31 (4.38)

[a] Mean [b] Standard Deviation

Copas Rate: Given the findings reported above and in previous chapters in relation to age at index offence, previous court appearances at which convicted, and age at first conviction, it was decided to calculate the Copas Rate (Copas & Marshall, 1998; Francis, Harman, & Humphreys, 2005) and to test for group differences. The Copas Rate forms part of the OGRS2 algorithm and uses the time since the offender's first conviction and the number of court appearances at which convicted to calculate a rate of conviction. The equation used to calculate the Copas Rate is:

$$75\sqrt{\frac{x}{(y+5)}}$$

where *x* is the total number of court appearances at which convicted and *y* is the number of years since the offender's first conviction.

Table 5.7 displays the mean Copas Rate by group. The Copas Rate was calculated within SPSS and the Kolmogorov-Smirnov tests were ran to test for normality. No problems with normality were observed and the assumption of homogeneity of variances was also met (Levene statistic = 0.319, $p = 0.727$). A one-way ANOVA was undertaken to test for differences between the three groups on the rate of conviction. Significant differences were observed between the groups: F = 7.467, $df = 2, 292$, $p = 0.001$. Post hoc Scheffe tests indicated that there were significant differences between the completers and non-starters ($p = 0.001$) only. The non-starters had a significantly higher rate of conviction than the completers.

Table 5.7: Mean Copas Rate by group.

	Copas Rate
Completers	51.28[a] (18.45[b])
Non-completers	58.48 (18.39)
Non-starters	60.09 (18.65)
Total	56.07 (18.93)

[a] Mean [b] Standard Deviation

Previous custody (under/over 21/total): Table 5.8 presents the mean number of custodial sentence by group. As with the age of first conviction, the Kolmogorov-Smirnov test revealed that these data were non-normal (Custody under 21: completers: *K-S Z* = 3.972, $p < 0.001$, non-completers: *K-S Z* = 2.096, $p < 0.001$, non-starters: *K-S Z* = 3.405, $p < 0.001$; Custody 21 and above: completers: *K-S Z* = 4.049, $p < 0.001$, non-completers: *K-S Z* = 2.913, $p < 0.001$, non-starters: *K-S Z* = 4.106, $p < 0.001$; Total custody: completers: *K-S Z* = 3.219, $p < 0.001$, non-completers: *K-S Z* = 2.047, $p < 0.001$, non-starters: *K-S Z* = 2.975, $p < 0.001$). Three Kruskal-Wallis tests were therefore undertaken; no significant differences between the groups were found on any of the three variables: Custody under 21: χ^2 (2, *N* = 293) = 3.372, $p = 0.185$, $\Phi_c = 0.11$, Custody 21 and above, χ^2 (2, *N* = 293) = 2.224, $p = 0.329$, $\Phi_c = 0.09$ and Total custody, χ^2 (2, *N* = 293) = 1.664, $p = 0.435$, $\Phi_c = 0.07$.

Table 5.8: Mean number of custodial sentences by group

	Custody under 21	Custody 21 and above	Total custody
Completers	0.99[a] (2.48[b])	0.88 (1.66)	1.87 (3.33)
Non-completers	0.98 (1.56)	1.04 (3.04)	2.02 (3.70)
Non-starters	1.08 (1.80)	1.37 (3.02)	2.45 (3.86)
Total	1.03 (2.08)	1.11 (2.53)	2.14 (3.66)

[a] Mean [b] Standard Deviation

Number of different offence categories: Table 5.9 presents the mean number of different offence categories for which each group have been convicted.

Table 5.9: Mean number of offence categories for which convicted by group.

	Offence categories
Completers	3.55[a] (1.96[b])
Non-completers	3.80 (2.07)
Non-starters	3.96 (1.86)
Total	3.76 (1.94)

[a] Mean [b] Standard Deviation

Once again, these data violated the normality assumption: completers: *K-S Z* = 2.089, $p < 0.001$, non-starters: *K-S Z* = 1.532, $p = 0.018$, hence the non-parametric Kruskal-Wallis test was used to test for group differences. No significant differences were found: χ^2 (2, $N = 292$) = 3.631, $p = 0.163$, $\Phi_c = 0.11$.

Previous violent offences: The mean number of previous violent convictions by group is displayed in table 5.10. The distribution of these offences was non-normal (completers: *K-S Z* = 2.861, $p < 0.001$, non-completers: *K-S Z* = 1.641, $p = 0.000$, non-starters: *K-S Z* = 2.685, $p < 0.001$) and hence required non-parametric tests to assess for significant differences between the groups. A Kruskal-Wallis test found no significant difference between the groups in relation to their number of previous violence offences: χ^2 (2, $N = 293$) = 1.794, $p = 0.408$, $\Phi_c = 0.08$.

Table 5.10: Mean number of previous violent and sexual convictions by group

	Violent convictions	Sexual convictions
Completers	1.26[a] (1.62[b])	0.13 (0.84)
Non-completers	2.00 (2.89)	0 (0)
Non-starters	1.45 (2.04)	0.06 (0.27)
Total	1.45 (2.04)	0.08 (0.58)

[a] Mean [b] Standard Deviation

Previous sex offences: The mean number of previous sexual convictions by group is displayed in table 5.10. As can be seen, very low numbers of the sample were previously convicted of sexual offences. As with the violent convictions, the data was non-normal (completers: *K-S Z* = 5.664, $p < 0.001$, non-starters: *K-S Z* = 5.921, $p < 0.001$) and hence a Kruskal-Wallis test was employed to test for significant differences between the groups. This test failed to find any significant differences: χ^2 (2, *N* = 293) = 2.525, $p = 0.283$, $\Phi_c = 0.09$.

Need variables:

OASys data: Total OASys score: Table 5.11 presents the groups' mean OASys total scores. As can be seen, the non-starters' mean score seems to be particularly higher than the completers and non-completers. As the Kolmogorov-Smirnov test indicated no problems with normality (completers: *K-S Z* = 0.846, $p < 0.472$, non-completers: *K-S Z* = 0.696, $p = 0.709$, non-starters: *K-S Z* = 0.584, $p < 0.885$) and the Levene's statistic indicates no problems with homogeneity of variance (0.277, $p = 0.758$), a one-way ANOVA was conducted on the data to test for differences between the groups. A significant effect was observed: F = 7.588, *df* = 2, 292, $p = 0.001$. Post hoc Scheffe tests indicated that there were significant differences between the completers' and non-starters ($p = 0.001$) and the non-completers' and the non-starters' ($p = 0.034$) OASys total scores. The non-starters group hence has significantly higher OASys total scores than the other two groups.

Table 5.11: Mean OASys total score by group.

	OASys total score
Completers	66.79[a] (26.85[b])
Non-completers	66.98 (28.04)
Non-starters	79.29 (26.70)
Total	72.02 (27.57)

[a] Mean [b] Standard Deviation

OASys data: Total OASys criminogenic and non-criminogenic scores: Table 5.12 displays the groups' total OASys criminogenic and non-criminogenic need scores. Both variables met the assumptions of normality and homogeneity of variances (criminogenic score: completers: *K-S Z* = 1.354, *p* < 0.051, non-completers: *K-S Z* = 0.818, *p* = 0.515, non-starters: *K-S Z* = 1.237, *p* < 0.094 and Levene's statistic = 2.514, *p* = 0.08; non-criminogenic score: completers: *K-S Z* = 1.130, *p* < 0.155, non-completers: *K-S Z* = 0.565, *p* = 0.907, non-starters: *K-S Z* = 1.146, *p* < 0.145 and Levene's statistic = 0.314, *p* = 0.73).

Table 5.12: Mean OASys criminogenic and non-criminogenic needs scores by group.

	Criminogenic score	Non-criminogenic score
Completers	29.35[a] (19.84[b])	14.06 (10.84)
Non-completers	28.69 (16.88)	15.47 (10.26)
Non-starters	36.54 (21.59)	14.13 (9.77)
Total	32.24 (20.44)	14.30 (10.29)

Two one-way ANOVAs were therefore performed on the data to test for significant differences between the three groups. In relation to the criminogenic need variable, a significant difference was found between the groups, F = 4.758, *df* = 2, 293, *p* = 0.009. Post-hoc Scheffe tests revealed that the completers and non-starters had

significantly different scores (p = 0.021) with non-starters having a significantly higher criminogenic needs score. No other comparisons were significant.

In relation to the non-criminogenic need variable, no significant differences were observed between the groups, F = 0.339, df = 2, 293, p = 0.712.

OASys data: Individual need scores: Table 5.13 displays the groups' mean individual OASys need scores. Kolmogorov-Smirnov tests indicated that all of these individual need scores violated the assumption of normality: see table 5.14.

Table 5.13: OASys individual need scores by group

	Completers	Non-completers	Non-starters	Total	χ^2 (*p*)	Φ_c
Criminal history (sections 1 and 2)	23.32[a] (14.78[b])	22.62 (13.81)	28.67 (14.06)	25.44 (14.55)	10.501 (0.005)	0.19
Accommodation	2.87 (4.21)	2.76 (4.29)	3.67 (4.62)	3.19 (4.04)	3.294 (0.193)	0.10
Financial Management and Income	10.07 (6.05)	10.58 (6.01)	12.04 (5.52)	10.97 (5.88)	6.733 (0.035)	0.15
Education, Training and Employability	3.74 (3.68)	3.38 (3.42)	4.39 (3.75)	3.96 (3.68)	3.859 (0.145)	0.11
Relationships	1.82 (1.41)	1.60 (1.29)	1.77 (1.30)	1.76 (1.34)	0.751 (0.687)	0.05
Lifestyle and Associates	6.23 (5.04)	7.31 (5.26)	7.82 (5.06)	7.06 (5.12)	6.151 (0.046)	0.14
Drug Misuse	2.21 (3.90)	2.60 (3.59)	3.06 (4.27)	2.62 (4.02)	4.998 (0.082)	0.13
Alcohol Misuse	1.81 (1.68)	1.73 (1.78)	2.05 (1.81)	1.90 (1.75)	1.270 (0.530)	0.07
Emotional Wellbeing	2.16 (2.20)	1.60 (2.25)	1.98 (2.10)	2.00 (2.17)	3.575 (0.167)	0.11
Thinking and Behaviour	8.83 (2.77)	8.56 (2.93)	8.98 (2.63)	8.85 (2.74)	1.080 (0.583)	0.06
Attitudes	3.72 (3.23)	4.22 (4.04)	4.91 (3.54)	4.30 (3.52)	7.746 (0.021)	0.16

Table 5.14: Kolmogorov-Smirnov output for the OASys individual need scores

	Completers	Non-completers	Non-starters
Criminal history (sections 1 and 2)	1.620 (p = 0.010)	1.026 (p = 0.243)	1.828 (p = 0.03)
Accommodation	3.188 (p < 0.001)	2.427 (p < 0.001)	2.711 (p < 0.001)
Financial Management and Income	2.740 (p < 0.001)	1.447 (p = 0.030)	2.683 (p < 0.001)
Education, Training and Employability	2.482 (p < 0.001)	1.422 (p = 0.035)	2.178 (p < 0.001)
Relationships	2.189 (p < 0.001)	1.577 (p = 0.014)	2.471 (p < 0.001)
Lifestyle and Associates	2.330 (p < 0.001)	1.534 (p = 0.018)	2.203 (p < 0.001)
Drug Misuse	4.010 (p < 0.001)	1.715 (p = 0.006)	3.179 (p < 0.001)
Alcohol Misuse	2.605 (p < 0.001)	1.579 (p = 0.014)	2.557 (p < 0.001)
Emotional Wellbeing	2.522 (p < 0.001)	2.276 (p < 0.001)	2.533 (p < 0.001)
Thinking and Behaviour	3.004 (p < 0.001)	1.787 (p = 0.003)	2.739 (p < 0.001)
Attitudes	2.387 (p < 0.001)	1.774 (p = 0.004)	2.024 (p = 0.001)

A series of Kruskal-Wallis tests were therefore performed on these data. Due to the large number of comparisons, a Bonferroni adjustment was made to the familywise error rate: 0.05/11 = 0.0045. None of the comparisons returned any statistically significant differences between the groups (see table 5.13), despite the effect sizes indicating moderate associations between group and scores in sections 1/2 (criminal history),

OASys data: Individual criminogenic and non-criminogenic need scores: Table 5.15 presents the mean individual criminogenic scores broken down by group. A series of Kolmogorov-Smirnov tests indicated that the distributions of these variables were not normal (see table 5.16), hence non-parametric Kruskal-Walls was used to test for group differences in the scores obtained. Given the number of comparisons undertaken, a Bonferonni adjustment was made to the familywise error rate (Dancey & Reidy, 2004): 0.05/10 = 0.005. None of the comparisons reached this level of significance (see table 5.15).

Table 5.15: OASys individual criminogenic need scores by group

	Completers	Non-completers	Non-starters	Total	χ^2 (*p*)	Φ_c
Accommodation	1.46 (3.57)	0.84 (2.47)	1.98 (4.11)	1.58 (3.68)	2.513 (0.285)	0.09
Financial Management and Income	3.93 (6.77)	2.73 (5.74)	5.80 (7.53)	4.52 (7.03)	7.829 (0.020)	0.16
Education, Training and Employability	2.21 (3.89)	1.24 (2.73)	2.63 (4.03)	2.24 (3.81)	4.807 (0.090)	0.13
Relationships	0.90 (1.48)	0.78 (1.22)	1.04 (1.44)	0.94 (1.42)	1.296 (0.523)	0.07
Lifestyle and Associates	5.33 (5.53)	7.02 (5.42)	7.21 (5.47)	6.37 (5.55)	8.698 (0.014)	0.17
Drug Misuse	1.40 (3.47)	1.60 (3.50)	2.30 (4.22)	1.81 (3.81)	4.698 (0.095)	0.13
Alcohol Misuse	1.37 (1.76)	1.67 (1.82)	1.68 (1.94)	1.54 (1.85)	1.942 (0.379)	0.08
Emotional Wellbeing	1.36 (2.20)	0.82 (1.86)	1.39 (2.17)	1.29 (2.14)	3.021 (0.221)	0.10
Thinking and Behaviour	8.60 (3.10)	8.47 (3.13)	8.83 (2.92)	8.68 (3.02)	0.796 (0.672)	0.05
Attitudes	2.79 (3.56)	3.51 (4.34)	3.67 (4.05)	3.27 (3.90)	2.607 (0.272)	0.09

Table 5.16: Kolmogorov-Smirnov output for the OASys individual criminogenic need scores

	Completers	Non-completers	Non-starters
Accommodation	5.344 ($p < 0.001$)	3.358 ($p < 0.001$)	5.035 ($p < 0.001$)
Financial Management and Income	5.043 ($p < 0.001$)	3.241 ($p < 0.001$)	4.169 ($p < 0.001$)
Education, Training and Employability	4.747 ($p < 0.001$)	3.042 ($p < 0.001$)	3.951 ($p < 0.001$)
Relationships	4.362 ($p < 0.001$)	2.712 ($p < 0.001$)	4. 015 ($p < 0.001$)
Lifestyle and Associates	2.573 ($p < 0.001$)	1.473 ($p < 0.001$)	2.240 ($p < 0.001$)
Drug Misuse	5.320 ($p < 0.001$)	3.046 ($p < 0.001$)	4.556 ($p < 0.001$)
Alcohol Misuse	4.039 ($p < 0.001$)	1.922 ($p = 0.001$)	3.753 ($p < 0.001$)
Emotional Wellbeing	4.561 ($p < 0.001$)	3.157 ($p < 0.001$)	4.363($p < 0.001$)
Thinking and Behaviour	3.018 ($p < 0.001$)	1.782 ($p = 0.003$)	2.719 ($p < 0.001$)
Attitudes	3.098 ($p < 0.001$)	1.733 ($p = 0.005$)	2.512 ($p < 0.001$)

The above analysis was repeated for the non-criminogenic needs scores derived from the OASys data. The means score of these variables broken down by group can be seen in table 5.17. Again, there were problems with non-normality amongst the variables (see table 5.18 for the Kolmogorov-Smirnov values); hence non-parametric Kruskal-Wallis was used to test for differences between the groups. As with the criminogenic need variables, the Bonferroni adjustment was used. None of the comparisons reached the significance level of $p < 0.005$ (see table 5.17).

Table 5.17: OASys individual non-criminogenic need scores by group

	Completers	Non-completers	Non-starters	Total	χ^2 (*p*)	Φ_c
Accommodation	1.35 (3.00)	1.91 (3.95)	1.69 (3.35)	1.58 (3.68)	0.714 (0.700)	0.05
Financial Management and Income	6.14 (6.28)	7.84 (6.86)	6.25 (6.83)	6.45 (6.61)	2.698 (0.260)	0.10
Education, Training and Employability	1.52 (2.30)	2.13 (3.11)	1.76 (2.68)	1.72 (2.59)	0.394 (0.821)	0.04
Relationships	0.91 (1.21)	0.82 (1.21)	0.72 (1.08)	0.82 (1.15)	1.130 (0.568)	0.06
Lifestyle and Associates	0.90 (2.13)	0.29 (1.55)	0.61 (2.11)	0.69 (2.05)	7.836 (0.020)	0.16
Drug Misuse	0.81 (2.33)	0.96 (1.99)	0.75 (1.98)	0.81 (2.13)	2.575 (0.276)	0.09
Alcohol Misuse	0.44 (0.97)	0.07 (0.25)	0.37 (0.88)	0.35 (0.87)	5.113 (0.078)	0.13
Emotional Wellbeing	0.80 (1.49)	0.64 (1.51)	0.59 (1.16)	0.69 (1.36)	1.588 (0.452)	0.07
Thinking and Behaviour	0.22 (1.37)	0.09 (0.60)	0.16 (1.11)	0.17 (1.17)	0.180 (0.914)	0.02
Attitudes	0.94 (1.76)	0.71 (1.60)	1.24 (2.30)	1.03 (1.99)	1.202 (0.548)	0.06

Table 5.18: Kolmogorov-Smirnov output for the OASys individual non-criminogenic need scores

	Completers	Non-completers	Non-starters
Accommodation	4.530 ($p < 0.001$)	2.959 ($p < 0.001$)	4.211 ($p < 0.001$)
Financial Management and Income	2.436 ($p < 0.001$)	1.201 ($p = 0.112$)	2.988 ($p < 0.001$)
Education, Training and Employability	3.301 ($p < 0.001$)	2.073 ($p < 0.001$)	3.427 ($p < 0.001$)
Relationships	3.700 ($p < 0.001$)	2.059 ($p < 0.001$)	3.847 ($p < 0.001$)
Lifestyle and Associates	5.232 ($p < 0.001$)	3.553 ($p < 0.001$)	5.592 ($p < 0.001$)
Drug Misuse	5.175 ($p < 0.001$)	2.655 ($p < 0.001$)	5.169 ($p < 0.001$)
Alcohol Misuse	5.282 ($p < 0.001$)	3.606 ($p < 0.001$)	5.409 ($p < 0.001$)
Emotional Wellbeing	4.441 ($p < 0.001$)	3.121 ($p < 0.001$)	4.499($p < 0.001$)
Thinking and Behaviour	5.982 ($p < 0.001$)	3.603 ($p < 0.001$)	5.865 ($p < 0.001$)
Attitudes	3.893 ($p < 0.001$)	2.565 ($p < 0.001$)	3.888 ($p < 0.001$)

OASys: Risk classification: Table 5.19 presents the groups' OASys risk classifications. This association between group and risk level was found to be significant: χ^2 (4, *N* = 293) = 12.587, *p* = 0.013, Φ_c = 0.15. Analysis of the distribution shows that there are fewer non-starters than non-completers or completers within the low risk classification and consequently more within the high risk classification.

Table 5.19: Distribution of OASys risk classifications by group

	Low	Medium	High	Total
Completers (%)	23 (18.3)	88 (69.8)	15 (11.9)	126 (100)
Non-completers (%)	7 (15.6)	31 (68.9)	7 (15.6)	45 (100)
Non-starters (%)	7 (5.7)	87 (71.3)	28 (23.0)	122 (100)
Total (%)	37 (12.6)	206 (70.3)	50 (17.1)	293 (100)

Prediction of group from variables which significantly vary between groups:
A multinomial logistic regression was performed to determine whether it was possible to predict group from the variables that had differed significantly or significantly associated with group membership within the univariate analyses. These variables were previous breach, number of court appearances at which convicted before the age of 18, age at first conviction, Copas Rate, OASys total score and OASys risk classification. It was not possible, however, to enter all of these variables into one regression analysis as to do so would violate the assumption of independence amongst the predictor variables. As such, the OASys criminogenic needs score and the OASys risk classifications variables were not entered to ensure independence of the predictor variable, OASys total score. This variable was chosen over the OASys risk classification variable as, of the two, it was the continuous variable and hence retained variance that the bandings loses. The analysis will be repeated with the OASys criminogenic needs variable and the analyses compared to see which model fits the data better.

Additionally, it was thought unwise to enter the Copas Rate variable alongside the number of court appearances at which convicted before the age of 18 or the age at first conviction as these variables contribute to the calculation of the Copas Rate. As

such, three regressions were run. The first contained the previous breach, number of court appearances at which convicted before the age of 18, age at first conviction, and the OASys total score. The second substituted the OASys total score with the OASys criminogenic score. These analyses were then compared and the one that contributed the most variance in the model was used to form the final regression. This regression however, entered the Copas Rate variable as a substitute for the number of court appearances at which convicted before the age of 18 and the age at first conviction variables.

Regression one: For the analysis containing the OASys total score, there was a good model fit on the basis of the entered predictors, χ^2 (554, N = 293) = 550.629, p = 0.532, using a deviance criterion. Comparisons of the log-likelihood ratios for models with and without predictors showed a reliable improvement with the addition of the predictors (χ^2 (8, N = 293) = 25.020, p = 0.002). Correct classification rates were 64.3% for the Completers, 0% for the Non-completers condition, and 59.8% for the Non-starters condition. The overall correct classification was 52.6%, which represents an increase of 19.3% over by-chance accuracy. Analysis of the contribution of individual predictors to the model with and without each predictor revealed that the only predictor to significantly predict the outcome was OASys total score (χ^2 (2, N = 293) = 7.245, p = 0.027). All other predictors were non-significant (Previous breach: χ^2 (2, N = 293) = 1.338, p = 0.512; Number of court appearances at which convicted under 18 years: χ^2 (2, N = 293) = 2.595, p = 0.273; Age at first conviction: χ^2 (2, N = 293) = 2.873, p = 0.238). Analysis of the odds ratios (see table 5.20) indicates that with each unit increase in OASys total score the odds of being in the non-completer group compared to the non-starter group decrease by 3% (p = 0.009).

Table 5.20: *Multinomial logistic regression of programme completion as a function of previous breach, OASys total score, number of court appearances at which convicted under 18 and age of first conviction.*

		B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Non-completers vs. Completers (ref)	Intercept	0.927	1.260	0.541		
	Convictions < 18	0.114	0.070	2.623	1.121	0.976 – 1.287
	Age at first conviction	-0.065	0.053	1.508	0.937	0.884 – 1.040
	OASys total	0.014	0.009	2.637	0.986	0.970 – 1.003
	Previous breach = 0	-0.356	0.405	0.772	0.700	0.316 – 1.550
Non-starters vs. Completers (ref)	Intercept	0.279	0.965	0.084		
	Convictions < 18	0.048	0.055	0.734	1.049	0.941 – 1.169
	Age at first conviction	-0.054	0.040	1.824	0.948	0.877 – 1.025
	OASys total	0.008	0.006	1.820	1.008	0.996 – 1.021
	Previous breach = 0	-0.307	0.305	1.011	0.736	0.404 – 1.338
Non-completers vs. Non-starters (ref)	Intercept	0.648	1.304	0.247		
	Convictions < 18	0.066	0.064	1.095	1.069	0.944 – 1.210
	Age at first conviction	-0.012	0.056	0.044	0.988	0.885 – 1.104
	OASys total	-0.022	0.009	6.741	0.978**	0.961 – 0.995
	Previous breach = 0	-0.049	0.417	0.014	0.952	0.421 – 2.154

** $p < 0.01$

Regression two: The analysis containing the OASys criminogenic need score also returned a good model fit on the basis of the entered predictors, χ^2 (554, N = 293) = 542.900, *p* = 0.624, using a deviance criterion. Comparisons of the log-likelihood ratios for models with and without predictors showed a reliable improvement with the addition of the predictors (χ^2 (8, N = 293) = 23.385, *p* = 0.003). Correct classification rates were 62.7% for the Completers, 0% for the Non-completers condition, and 59.8% for the Non-starters condition. The overall correct classification was 51.9%, which represents an increase of 18.6% over by-chance accuracy. Analysis of the contribution of individual predictors to the model with and without each predictor revealed that the none of the predictors significantly predicted the outcome: OASys criminogenic score (χ^2 (2, N = 293) = 5.610, *p* = 0.061), Previous breach: χ^2 (2, N = 293) = 2.167, *p* = 0.338; Number of court appearances at which convicted under 18 years: χ^2 (2, N = 293) = 1.888, *p* = 0.389; Age at first conviction: χ^2 (2, N = 293) = 3.398, *p* = 0.183). The odds ratios are displayed in table 5.21.

Table 5.21: *Multinomial logistic regression of programme completion as a function of previous breach, OASys criminogenic score, number of court appearances at which convicted under 18 and age of first conviction.*

		B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Non-	Intercept	- 0.135	1.023	0.017		
completers vs.	Convictions < 18	0.092	0.069	1.795	1.096	0.958 – 1.254
Completers	Age at first conviction	-0.046	0.051	0.824	0.955	0.864 – 1.055
(ref)	OASys criminogenic	-0.008	0.010	0.727	0.992	0.972 – 1.011
	Previous breach = 0	-0.138	0.377	0.134	0.871	0.416 – 1.823
Non-starters vs.	Intercept	0.785	0.799	0.966		
Completers	Convictions < 18	0.049	0.055	0.833	1.051	0.945 – 1.168
(ref)	Age at first conviction	-0.067	0.040	2.839	0.935	0.865 – 1.011
	OASys criminogenic	0.012	0.007	2.891	1.012	0.998 – 1.025
	Previous breach = 0	-0.416	0.284	2.138	0.660	0.378 – 1.152
Non-	Intercept	-0.920	1.066	0.388		
completers vs.	Convictions < 18	0.043	0.062	0.469	1.043	0.924 – 1.179
Non-starters	Age at first conviction	0.021	0.055	0.145	1.021	0.917 – 1.137
(ref)	OASys criminogenic	-0.020	0.010	4.186	0.980*	0.962 – 0.999
	Previous breach = 0	0.278	0.388	0.512	1.320	0.617 – 2.825

** $p < 0.01$

Regression three: In evaluating the two regression models above, it was determined that the first regression (that containing the total OASys as opposed to the OASys criminogenic score) accounted for the most variance in the data: Regression one Nagelkerke = 0.94 compared with Regression two Nagelkerke = 0.88. Therefore the third regression attempted to predict group membership using the previous breach, Copas Rate (instead of the court appearances at which conviction under age 18 and age at first conviction) and OASys total score (as opposed to OASys criminogenic score) variables. This would establish whether the model containing the Copas Rate was more predictive than that containing the court appearances at which convicted under age 18 and the age at first conviction variables.

The regression analysis returned a good model fit on the basis of the entered predictors, χ^2 (566, N = 293) = 562.131, p = 0.538, using a deviance criterion. Comparisons of the log-likelihood ratios for models with and without predictors showed a reliable improvement with the addition of the predictors (χ^2 (6, N = 293) = 23.562, p = 0.001). Correct classification rates were 63.5% for the Completers, 0% for the Non-completers condition, and 55.7% for the Non-starters condition. The overall correct classification was 50.5%, which represents an increase of 17.2% over by-chance accuracy. Analysis of the contribution of individual predictors to the model with and without each predictor revealed that two of the predictors significantly predicted the outcome: OASys total score (χ^2 (2, N = 293) = 7.543, p = 0.023), Copas Rate χ^2 (2, N = 293) = 7.313, p = 0.026. The Previous Breach variable did not predict the outcome: χ^2 (2, N = 293) = 0.167, p = 0.920. Analysis of the odds ratios (see table 5.22) indicates that with each unit increase in OASys total score the odds of being in the non-completer group compared to the non-starter group decrease by 2% (p = 0.011) and with each unit increase in Copas Rate the odds of being in the non-completer group compared to the completer group increase by 3% (p = 0.015)

Table 5.22: *Multinomial logistic regression of programme completion as a function of previous breach, OASys total score, and Copas Rate.*

		B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Non-completers vs. Completers (ref)	Intercept	-1.875	0.790	5.627		
	Copas Rate	0.029	0.012	5.902*	1.030	1.006 – 1.055
	OASys total	-0.011	0.008	1.755	0.989	0.973 – 1.005
	Previous breach = 0	-0.088	0.420	0.044	0.916	0.402 – 2.087
Non-starters vs. Completers (ref)	Intercept	-1.675	0.790	7.649		
	Copas Rate	0.017	0.009	3.574	1.017	0.999 – 1.034
	OASys total	0.010	0.006	2.991	1.010	0.999 – 1.022
	Previous breach = 0	-0.127	0.317	0.160	0.881	0.473 – 1.640
Non-completers vs. Non-starters (ref)	Intercept	-0.221	0.803	0.075		
	Copas Rate	0.013	0.012	1.153	1.013	0.989 – 1.037
	OASys total	-0.021	0.008	6.494*	0.979	0.963 – 0.995
	Previous breach = 0	0.039	0.435	0.008	1.040	0.443 – 2.439

** p < 0.05 ** p < 0.01

Starter vs. Non-starter analysis:

Given the indications within the preceding analyses that the non-starters differ from the non-completers and completers on a number of variables, the analyses were re-run to determine whether it was possible to predict non-starting a programme from the aforementioned variables. As such, the completer and non-completer groups were combined to form the 'starters' group and this group was compared to the non-starter group. Initially, univariate analyses were undertaken to determine which variables should be entered into the logistic regression.

Index offence: The distribution of offence categories across starter and non-starter groups is displayed in table 5.23.

Table 5.23: Distribution of offence categories of index offences across starter and non-starter groups (%).

Offence category	Starters	Non-starters
Motoring offences	41 (23.97)	37 (30.33)
Theft and handling stolen goods	37 (21.64)	28 (22.95)
Other (excluding motoring offences)	28 (16.37)	16 (13.11)
Violence against the person	32 (18.71)	13 (10.66)
Drug offences	11 (6.43)	6 (4.92)
Burglary	12 (7.02)	11 (9.02)
Fraud and forgery	7 (4.09)	4 (3.28)
Criminal damage	2 (1.17)	7(5.74)
Robbery	1 (0.58)	0 (0)
Sexual offences	0 (0)	0 (0)
Total	171 (100)	122 (100)

A chi-square analysis of this distribution (without the sexual offences row as no data were present) returned four cells (40.7%) with expected counts of less than five. As

with the earlier analysis, this meant that the use of chi-square on these data was unreliable as more than 20% of cells had expected counts below five (Tabachnick & Fidell, 2006). To resolve this, the collapsed classification structure as described earlier was used on these data. This distribution produced only one cell (8.3%) with an expected count less than five and hence it was deemed reliable to undertake the chi-square analysis with this distribution. Table 5.24 presents this distribution.

Table 5.24: *Distribution of offence categories of index offences (with the collapsed category of acquisitive crime) across starter and non-starter groups (%).*

Offence category	Starter	Non-starters
Acquisitive crime	57 (33.33)	43 (35.25)
Motoring offences	41 (23.98)	37 (30.33)
Other (excluding motoring offences)	28 (16.37)	16 (13.11)
Violence against the person	32 (18.71)	13 (10.66)
Drug offences	11 (6.43)	6 (4.92)
Criminal damage	2 (1.17)	7(5.74)
Total	171 (100)	122 (100)

This chi-square was non-significant: χ^2 (5, N = 293) = 9.788, p = 0.08, Φ_c = 0.18). Thus, there is no evidence of between-group differences in index offence type.

Previous breach of probation/parole/license/bail/community based sentence: Table 5.25 shows the number of offenders within the sample who have previously breached a criminal justice order (probation/parole/license/bail/community based sentence). Non-starters have the higher rate of breach with over two-thirds, compared with just over half of starters, previously having been breached. The association between group and breach was significant (χ^2 (2, N = 293) = 5.755, p = 0.016, Φ_c = 0.14).

Table 5.25: Breach by starter and non-starter group

	Yes	No	Total
Starters (%)	97 (56.7)	74 (43.3)	171 (100)
Non-starters (%)	86 (70.5)	36 (29.5)	122 (100)
Total (%)	183 (62.5)	110 (37.5)	293 (100)

Previous convictions (under/over 18/total): Table 5.26 presents details relating to the mean numbers of previous court appearances at which convicted, below the age of 18 years, from 18 years onwards and in total broken down by starter and non-starter group. Kolmogorov-Smirnov tests revealed that these data were non-normal (Convictions under 18: starters: *K-S Z* = 3.040, $p < 0.001$; non-starters: *K-S Z* = 2.104, $p < 0.001$; Convictions 18 and above: starters: *K-S Z* = 2.738, $p < 0.001$; non-starters: *K-S Z* = 2.247, $p < 0.001$; Total convictions: starters: *K-S Z* = 2.332, $p = 0.001$; non-starters: *K-S Z* = 1.605, $p = 0.012$). Three Mann Whitney U tests were therefore undertaken to test for significant differences between the starter and non-starter groups. Significant differences were observed in relation to the Convictions under 18 variable, $U = 8538.00$, $p = 0.006$ and the Total convictions variable, $U = 8831.00$, $p = 0.025$. Non-starters had significantly more convictions under the age of 18 and total convictions than the programme starters. The comparison of starters' and non-starters' number of convictions above 18 did not differ significantly, $U = 10304.5$, $p = 0.859$.

Table 5.26: Mean number of convictions by group

	Convictions under 18	Convictions 18 and above	Total convictions
Starters	2.15 (2.94)	5.50 (6.02)	7.65 (6.97)
Non-starters	3.16 (3.60)	5.98 (6.69)	9.13 (7.10)
Total	2.57 (3.27)	5.70 (6.30)	8.27 (7.05)

[a] Mean [b] Standard Deviation

Age at first conviction: Table 5.27 presents the mean ages of first conviction by starter and non-starter group. Kolmogorov-Smirnov tests revealed that these data were non-normal (starters: *K-S Z* = 2.326, *p* = 0.001, non-starters: *K-S Z* = 1.850, *p* = 0.002). A Mann Whitney U test returned a significant result, *U* = 8464.00, *p* = 0.006. Starters were significantly older than the non-starters at their first conviction.

Table 5.27: Mean age of first conviction by starter and non-starter group.

	Age at first conviction
Starters	17.88 (5.01)
Non-starters	16.51 (3.15)
Total	17.31 (4.38)

Copas Rate: As with the three group analysis above, the Copas Rate (Copas & Marshall, 1998) was calculated and to test for group differences. Table 5.28 displays the mean Copas Rate by group. No problems with normality were observed and the assumption of homogeneity of variances was also met (Levene's statistic = 0.290, p = 0.591). An independent samples t-test was undertaken to test for differences between the two groups on the rate of conviction. Significant differences were observed between the groups: *t* [291] = 3.129, *p* = 0.002. The non-starters had a significantly higher rate of conviction than the starters.

Table 5.28: Mean Copas Rate by group.

	Copas Rate
Starters	53.18 (18.65)
Non-starters	60.09 (18.65)
Total	56.07 (18.93)

[a] Mean [b] Standard Deviation

Previous custody (under/over 21/total): Table 5.29 presents the mean number of custodial sentences by starter and non-starter group. As with the age of first conviction, the Kolmogorov-Smirnov test revealed that these data were non-normal (Custody under 21: starters: *K-S Z* = 4.381, *p* < 0.001; non-starters: *K-S Z* = 3.405, *p* < 0.001; Custody 21 and above: starters: *K-S Z* = 4.779, *p* < 0.001; non-starters: *K-S Z* = 4.106, *p* < 0.001; Total custody: starters: *K-S Z* = 3.821, *p* < 0.001; non-starters: *K-S Z* = 2.975, *p* < 0.001). Three Mann Whitney U tests were therefore undertaken. No significant differences between the groups were found on any of the three variables: Custody under 21: *U* = 9579.50, *p* = 0.168, Custody 21 and above, *U* = 10207.00, *p* = 0.700 and Total custody, *U* = 9567.50, *p* = 0.198.

Table 5.29: Mean number of custodial sentences by group

	Custody under 21	Custody 21 and above	Total custody
Starters	0.99 (2.27)	0.92 (2.11)	1.91 (3.50)
Non-starters	1.08 (1.80)	1.37 (3.02)	2.45 (3.86)
Total	1.03 (2.08)	1.11 (2.53)	2.14 (3.66)

[a] Mean [b] Standard Deviation

Number of different offence categories: Table 5.30 presents the mean number of different offence categories for which each group has been convicted.

Table 5.30: Mean number of offence categories for which convicted by starter and non-starter group.

	Offence categories
Starters	3.62 (1.99)
Non-starters	3.96 (1.86)
Total	3.76 (1.94)

[a] Mean [b] Standard Deviation

Once again, these data violated the normality assumption: starters: *K-S Z* = 2.434, *p* < 0.001, non-starters: *K-S Z* = 1.532, *p* = 0.018, hence the non-parametric Mann Whitney U test was used to test for group differences. No significant differences were found: *U* = 9085.00, *p* = 0.067.

Previous violent offences: The mean number of previous violent convictions by starter and non-starter group is displayed in table 5.31. The distribution of these offences was non-normal (starters: *K-S Z* = 2.047, *p* < 0.001; non-starters: *K-S Z* = 2.685, *p* < 0.001) and hence required non-parametric tests to assess for significant differences between the groups. A Mann Whitney U test found no significant difference between the groups in relation to their number of previous violent offences: *U* = 10346.00, *p* = 0.901.

Previous sex offences: The mean number of previous sexual convictions by group is displayed in table 5.31. As with the violent convictions, the data were non-normal (starters: *K-S Z* = 6.678, *p* < 0.001, non-starters: *K-S Z* = 5.921, *p* < 0.001) and hence a Mann Whitney U test was employed to test for significant differences between the groups. This test failed to find any significant differences: *U* = 10348.00, *p* = 0.745.

Table 5.31: Mean number of previous violent and sexual convictions by starter and non-starter group

	Violent convictions	Sexual convictions
Starter	1.46 (2.05)	0.09 (0.72)
Non-starters	1.45 (2.04)	0.06 (0.27)
Total	1.45 (2.04)	0.08 (0.58)

[a] Mean [b] Standard Deviation

Need variables:

OASys data: Total OASys score: Table 5.32 presents the starter and non-starter groups' mean OASys total scores. The Kolmogorov-Smirnov test indicated no problems

with normality (starters: *K-S Z* = 0.892, *p* < 0.404; non-starters: *K-S Z* = 0.584, *p* < 0.885) and so an independent samples t-test was conducted on the data to test for differences between the two groups. Levene's test for equality of variances indicated no problems with heterogeneity of variance (F = 0.272, p = 0.602) and hence equal variances were assumed. A significant difference between the groups was observed: *t* [291] = 3.902, *p* < 0.001. The non-starters group has significantly higher OASys total scores than the starter group.

Table 5.32: Mean OASys total score by starter and non-starter group.

	OASys total score
Starters	66.84 (27.08)
Non-starters	79.29 (26.70)
Total	72.02 (27.57)

[a] Mean [b] Standard Deviation

OASys data: Total OASys criminogenic and non-criminogenic scores: Table 5.33 displays the starter and non-starter groups' total OASys criminogenic and non-criminogenic need scores. The OASys non-criminogenic need variable met the assumptions of normality and homogeneity of variances (starters: *K-S Z* = 1.239, *p* = 0.093, non-starters: *K-S Z* = 1.146, *p* = 0.145 and Levene's statistic = 0.568, *p* = 0.452) and hence an independent samples t-test was conducted to assess for any significant differences between the two groups. No significant differences were observed: *t* [291] = -0.242, *p* = 0.809.

Table 5.33: Mean OASys criminogenic and non-criminogenic needs scores by starter and non-starter groups.

	Criminogenic score	Non-criminogenic score
Starters	29.18 (19.06)	14.43 (10.67)
Non-starters	36.54 (21.59)	14.13 (9.77)
Total	32.24 (20.44)	14.30 (10.29)

The OASys criminogenic need variable did not meet the assumption of normality (starters: *K-S Z* = 1.470, *p* = 0.027) and hence a Mann Whitney U was conducted on the data. Significant differences were found between the starters and non-starters in relation to their scores on the OASys criminogenic need variable: *U* = 8376.00, *p* = 0.004. Programme non-starters had significantly higher OASys criminogenic needs scores than the programme starters.

OASys data: Individual need scores: Table 5.34 displays the starter and non-starter groups' mean individual OASys need scores. Kolmogorov-Smirnov tests indicated that all of these individual need scores violated the assumption of normality: see table 5.35.

Table 5.34: OASys individual need scores by starter and non-starter group

	Starters	Non-starters	Total	*U* (*p*)	r[a]
Criminal history (sections 1 and 2)	23.13 (14.49)	28.67 (14.06)	25.44 (14.55)	8139.0 (0.001)	-0.13
Accommodation	2.84 (4.22)	3.67 (4.62)	3.19 (4.04)	9288.0 (0.084)	-0.07
Financial Management and Income	10.20 (6.03)	12.04 (5.52)	10.97 (5.88)	8641.5 (0.011)	-0.10
Education, Training and Employability	3.64 (3.61)	4.39 (3.75)	3.96 (3.68)	9109.0 (0.060)	-0.07
Relationships	1.76 (1.38)	1.77 (1.30)	1.76 (1.34)	10272.0 (0.819)	-0.01
Lifestyle and Associates	6.51 (5.10)	7.82 (5.06)	7.06 (5.12)	8892.0 (0.029)	-0.09
Drug Misuse	2.31 (3.81)	3.06 (4.27)	2.62 (4.02)	9397.5 (0.109)	-0.06
Alcohol Misuse	1.79 (1.70)	2.05 (1.81)	1.90 (1.75)	9664.5 (0.266)	-0.05
Emotional Wellbeing	2.01 (2.22)	1.98 (2.10)	2.00 (2.17)	10281.5 (0.828)	-0.01
Thinking and Behaviour	8.75 (2.80)	8.98 (2.63)	8.85 (2.74)	10050.0 (0.585)	-0.02
Attitudes	3.87 (3.46)	4.91 (3.54)	4.30 (3.52)	8480.0 (0.006)	-0.11

[a] Calculated using the formula: r = z / √(N), where z is taken from the Mann Whitney U output and N is the number of observations (Field, 2005).

Table 5.35: Kolmogorov-Smirnov output for the OASys individual need scores

	Starters	Non-starters
Criminal history (sections 1 and 2)	1.821 (p = 0.003)	1.828 (p = 0.03)
Accommodation	3.990 (p < 0.001)	2.711 (p < 0.001)
Financial Management and Income	3.099 (p < 0.001)	2.683 (p < 0.001)
Education, Training and Employability	2.869 (p < 0.001)	2.178 (p < 0.001)
Relationships	2.699 (p < 0.001)	2.471 (p < 0.001)
Lifestyle and Associates	2.482 (p < 0.001)	2.203 (p < 0.001)
Drug Misuse	4.239 (p < 0.001)	3.179 (p < 0.001)
Alcohol Misuse	3.055 (p < 0.001)	2.557 (p < 0.001)
Emotional Wellbeing	3.120 (p < 0.001)	2.533 (p < 0.001)
Thinking and Behaviour	3.504 (p < 0.001)	2.739 (p < 0.001)
Attitudes	2.953 (p < 0.001)	2.024 (p = 0.001)

A series of Mann Whitney U tests were therefore performed on these data. Due to the large number of comparisons, a Bonferroni adjustment was made to the familywise error rate: 0.05/11 = 0.0045. Only one comparison returned a statistically significant difference between the groups (see table 5.34): sections 1/2 (criminal history), U = 8139.00, p < 0.001.

OASys data: Individual criminogenic and non-criminogenic need scores: Table 5.37 presents the mean individual criminogenic scores broken down by starter and non-starter group. Kolmogorov-Smirnov tests indicated that the distributions of these variables were non normally distributed (see table 5.36), hence non-parametric Mann Whitney U was used to test for group differences in the scores obtained. Given the number of comparisons undertaken, a Bonferonni adjustment was made to the

familywise error rate (Dancey & Reidy, 2004): 0.05/10 = 0.005. None of the comparisons reached this level of significance (see table 5.37).

Table 5.36: Kolmogorov-Smirnov output for the OASys individual criminogenic need scores

	Starters	Non-starters
Accommodation	6.309 ($p < 0.001$)	5.035 ($p < 0.001$)
Financial Management and Income	6.000 ($p < 0.001$)	4.169 ($p < 0.001$)
Education, Training and Employability	5.625 ($p < 0.001$)	3.951 ($p < 0.001$)
Relationships	5.128 ($p < 0.001$)	4. 015 ($p < 0.001$)
Lifestyle and Associates	2.799 ($p < 0.001$)	2.240 ($p < 0.001$)
Drug Misuse	6.136 ($p < 0.001$)	4.556 ($p < 0.001$)
Alcohol Misuse	4.462 ($p < 0.001$)	3.753 ($p < 0.001$)
Emotional Wellbeing	5.550 ($p < 0.001$)	4.363($p < 0.001$)
Thinking and Behaviour	3.512 ($p < 0.001$)	2.719 ($p < 0.001$)
Attitudes	3.536 ($p < 0.001$)	2.512 ($p < 0.001$)

Table 5.37: OASys individual criminogenic need scores by starter and non-starter group

	Starters	Non-starters	Total	*U* (*p*)	r[a]
Accommodation	1.30 (3.32)	1.98 (4.11)	1.58 (3.68)	9750.0 (0.168)	-0.06
Financial Management and Income	3.61 (6.52)	5.80 (7.53)	4.52 (7.03)	8893.0 (0.009)	-0.11
Education, Training and Employability	1.96 (3.64)	2.63 (4.03)	2.24 (3.81)	9327.0 (0.062)	-0.08
Relationships	0.87 (1.41)	1.04 (1.44)	0.94 (1.42)	9754.5 (0.271)	-0.04
Lifestyle and Associates	5.77 (5.54)	7.21 (5.47)	6.37 (5.55)	8839.0 (0.023)	-0.09
Drug Misuse	1.45 (3.47)	2.30 (4.22)	1.81 (3.81)	9308.5 (0.035)	-0.09
Alcohol Misuse	1.44 (1.78)	1.68 (1.94)	1.54 (1.85)	9861.0 (0.382)	-0.04
Emotional Wellbeing	1.22 (2.12)	1.39 (2.17)	1.29 (2.14)	9941.5 (0.405)	-0.03
Thinking and Behaviour	8.57 (3.10)	8.83 (2.92)	8.68 (3.02)	10042.5 (0.578)	-0.02
Attitudes	2.98 (3.78)	3.67 (4.05)	3.27 (3.90)	9.468.0 (0.155)	-0.06

[a] Calculated using the formula: r = z / √(N), where z is taken from the Mann Whitney U output and N is the number of observations (Field, 2005).

The above analysis was repeated for the non-criminogenic needs scores derived from the OASys data. The means score of these variables broken down by starter and non-starter group can be seen in table 5.39. Again, there were problems with non-normality amongst the variables (see table 5.38 for the Kolmogorov-Smirnov values); hence non-parametric Mann Whitney U was used to test for differences between the groups. As with the criminogenic need variables, the Bonferroni adjustment was used. None of the comparisons reached the significance level of $p < 0.005$ (see table 5.39).

Table 5.38: Kolmogorov-Smirnov output for the OASys individual non-criminogenic need scores

	Starters	Non-starters
Accommodation	5.401 ($p < 0.001$)	4.211 ($p < 0.001$)
Financial Management and Income	2.687 ($p < 0.001$)	2.988 ($p < 0.001$)
Education, Training and Employability	3.797 ($p < 0.001$)	3.427 ($p < 0.001$)
Relationships	4.240 ($p < 0.001$)	3.847 ($p < 0.001$)
Lifestyle and Associates	6.361 ($p < 0.001$)	5.592 ($p < 0.001$)
Drug Misuse	5.787 ($p < 0.001$)	5.169 ($p < 0.001$)
Alcohol Misuse	6.367 ($p < 0.001$)	5.409 ($p < 0.001$)
Emotional Wellbeing	5.419 ($p < 0.001$)	4.499($p < 0.001$)
Thinking and Behaviour	6.958 ($p < 0.001$)	5.865 ($p < 0.001$)
Attitudes	4.666 ($p < 0.001$)	3.888 ($p < 0.001$)

Table 5.39: OASys individual non-criminogenic need scores by starter and non-starter groups

	Starters	Non-starters	Total	*U* (*p*)	r[a]
Accommodation	1.50 (3.28)	1.69 (3.35)	1.58 (3.68)	9952.0 (0.398)	-0.03
Financial Management and Income	6.59 (6.46)	6.25 (6.83)	4.52 (7.03)	9963.5 (0.497)	-0.03
Education, Training and Employability	1.68 (2.54)	1.76 (2.68)	2.24 (3.81)	10348.0 (0.899)	-0.005
Relationships	0.72 (1.07)	0.72 (1.08)	0.94 (1.42)	9875.5 (0.313)	-0.04
Lifestyle and Associates	0.74 (2.01)	0.61 (2.11)	6.37 (5.55)	9920.5 (0.231)	-0.05
Drug Misuse	0.85 (2.24)	0.75 (1.98)	1.81 (3.81)	10215.5 (0.663)	-0.02
Alcohol Misuse	0.35 (0.86)	0.37 (0.88)	1.54 (1.85)	10386.5 (0.924)	-0.004
Emotional Wellbeing	0.76 (1.49)	0.59 (1.16)	1.29 (2.14)	10306.5 (0.826)	-0.01
Thinking and Behaviour	0.19 (1.21)	0.16 (1.11)	8.68 (3.02)	10382.0 (0.808)	-0.01
Attitudes	0.88 (1.71)	1.24 (2.30)	3.27 (3.90)	10019.0 (0.496)	-0.03

[a] Calculated using the formula: r = z / √(N), where z is taken from the Mann Whitney U output and N is the number of observations (Field, 2005).

OASys: Risk classification: Table 5.40 presents the starters and non-starter groups' OASys risk classifications. This association between group and risk level was found to be significant: χ^2 (2, *N* = 293) = 12.133, *p* = 0.002, Φ_c = 0.20. Analysis of the distribution shows that there are fewer non-starters than starters within the low risk classification and consequently more within the high risk classification.

Table 5.40: Distribution of OASys risk classifications by starter and non-starter groups

	Low	Medium	High	Total
Starters (%)	30 (17.5)	119 (69.6)	22 (12.9)	171 (100)
Non-starters (%)	7 (5.7)	87 (71.3)	28 (23.0)	122 (100)
Total (%)	37 (12.6)	206 (70.3)	50 (17.1)	293 (100)

Prediction of group from variables which significantly vary between starter and non-starter groups:

A logistic regression was performed to determine whether it was possible to predict programme starting from the previous breach, number of court appearances at which convicted before the age of 18, age at first conviction, and OASys total score variables. The total number of court appearances at which convicted was not included in the analysis due to issues of independence relating to the 'number of court appearances at which convicted before the age of 18' variable. The latter variable was chosen for entry above the former as the differences between groups on the total number of court appearances variable were accounted for by those court appearances at which convicted prior to the age of 18 years (there were no significant differences between the two groups on court appearances at which convicted over 18). Likewise, the OASys criminogenic needs score and the OASys risk classification variables were not entered to ensure independence of the predictor variable, OASys total score. This variable was chosen over the OASys risk classification variable as, of the two, it was the continuous variable and hence retained variance that the classification into bands loses. The analysis will be repeated with the OASys criminogenic needs variable and the analyses compared to see which model fits the data better. Additionally, the OASys individual needs score relating to Criminal History (sections one and two) was also not entered to

ensure the independence of the OASys total score variable. Finally, the Copas Rate was not entered initially due to the lack of independence from the number of court appearances at which convicted before the age of 18 and the age at first conviction variables. As with the three group analysis above, the two regressions (OASys total score versus OASys criminogenic score) were compared and that analysis accounting for the most variance in the dependent variable was re-run with the Copas Rate variable substituting the number of court appearances at which convicted before the age of 18 and the age at first conviction variables.

Regression one: For the analysis containing the OASys total score, there was a good model fit on as measured by the Hosmer and Lemeshow Test: χ^2 (8, *N* = 293) = 2.335, *p* = 0.969. This model was also significantly better than the constant-only model (block 0) containing only the intercept but no predictor variables: χ^2 (4, *N* = 293) = 17.046, *p* = 0.002. The model classified 60.1% of cases correctly which represents a 10.1% increase over chance: 33.6% of the non-starters and 78.9% of the starters were correctly classified. Table 5.41 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. Analysis of the contribution of individual predictors to the model revealed, as with the logistic regressions above, that the only predictor to significantly predict the outcome was OASys total score: B = -0.012, *p* = 0.034. All other predictors were non-significant (Previous breach: B = 0.215, *p* = 0.454; Number of court appearances at which convicted under 18 years: B = -0.007, *p* = 0.887; Age at first conviction: B = 0.041, *p* = 0.272). Analysis of the odds ratios (see table 5.41) indicates that with each unit increase in OASys total score the odds of being in the non-starter group compared to the starter group decrease by 1.2%.

Table 5.41: Logistic regression of programme starting as a function of previous breach, number of court appearances at which convicted under 18, age of first conviction, and OASys total score.

	B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Breaches (1)	0.215	0.287	0.561	1.240	0.707 – 2.175
Convictions < 18	-0.007	0.047	0.020	0.993	0.906 – 1.090
Age at first conviction	0.041	0.038	1.205	1.042	0.968 – 1.122
OASys total	-0.012	0.006	4.491*	0.988	0.977 – 0.999
Constant	0.467	0.901	0.269	0.1595	

* $p < 0.05$

Regression two: The analysis containing the OASys criminogenic need score also returned a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, *N* = 293) = 10.801, *p* = 0.213. This model was also significantly better than the constant-only model (block 0) containing only the intercept but no predictor variables: χ^2 (4, *N* = 293) = 17.348, *p* = 0.002. The model classified 62.5% of cases correctly which represents a 12.5% increase over chance: 33.6% of the non-starters and 83.0% of the starters were correctly classified. Table 5.42 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. Analysis of the contribution of individual predictors to the model revealed, as with the logistic regressions above, that the only predictor to significantly predict the outcome was OASys criminogenic need score: B = -0.014, *p* = 0.028. All other predictors were non-significant (Previous breach: B = 0.384, *p* = 0.151; Number of court appearances at which convicted under 18 years: B = -0.016, *p* = 0.736; Age at first conviction: B = 0.059, *p* = 0.117). Analysis of the odds ratios (see table 5.42) indicates that with each unit increase in OASys criminogenic need score the odds of being in the non-starter group compared to the starter group decrease by 1.4%.

Table 5.42: Logistic regression of programme starting as a function of previous breach, number of court appearances at which convicted under 18, age of first conviction, and OASys criminogenic score.

	B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Breaches (1)	0.384	0.268	2.061	1.469	0.869 – 2.481
Convictions < 18	-0.016	0.026	0.114	0.985	0.899 – 1.078
Age at first conviction	0.059	0.038	2.454	1.061	0.985 – 1.143
OASys criminogenic	-0.014	0.006	4.820*	0.986	0.974 – 0.999
Constant	-0.324	0.747	0.188	0.723	

* $p < 0.05$

Regression three: The above two regressions were compared using the Nagelkerke statistics to determine which model accounted for the most variance. The model containing the OASys criminogenic variable (as opposed to the OASys total score) had the highest Nagelkerke score (0.77 and 0.76 respectively). As such the logistic regression was re-run with the following predictors: Previous Breach, Copas Rate (in place of the total number of court appearances at which convicted and age at first conviction variables) and the OASys criminogenic need score. The analysis indicated a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, *N* = 293) = 4.425, *p* = 0.817. This model was also significantly better than the constant-only model (block 0) containing only the intercept but no predictor variables: χ^2 (3, *N* = 293) = 15.651, *p* = 0.001. The model classified 61.1% of cases correctly which represents a 11.1% increase over chance: 32.0% of the non-starters and 81.9% of the starters were correctly classified. Table 5.43 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. Analysis of the contribution of individual predictors to the model revealed that the only predictor to significantly predict the outcome was OASys criminogenic need score: B = -0.014, *p* = 0.030. All other predictors were non-significant (Previous breach: B = 0.271, *p* = 0.145; Copas Rate: B = 0.13, *p* = 0.097). Analysis of the odds ratios (see table 5.43) indicates that with each unit

increase in OASys criminogenic need score the odds of being in the non-starter group compared to the starter group decrease by 1.3%.

Table 5.43: Logistic regression of programme starting as a function of previous breach, Copas Rate, and OASys criminogenic score.

	B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Breaches (1)	0.271	0.287	0.890	1.311	0.747 – 2.301
Copas Rate	-0.13	0.008	2.759	0.988	0.973 – 1.002
OASys criminogenic	-0.014	0.006	4.725*	0.987	0.975 – 0.999
Constant	1.392	0.505	7.597	4.022	

* $p < 0.05$

Summary of findings: The table below summarises the findings reported within this chapter.

Table 5.44: Summary of significant findings.

Variable	Three group (completer, non-completer, non-starter	Two group (starter, non-starter)
Index offence type	No significance	No significance
Previous breach	Completers < Non-completers < Non-starters	Starters < Non-starters
Previous convictions < 18	Completers < Non-starters	Starters > Non-starters
Previous convictions > 18	No significance	No significance
Previous convictions total	No significance	Starters > Non-starters
Age at first conviction	Completers < Non-starters	Starters < Non-starters
Copas Rate	Completers < Non-starters	Starters > Non-starters
Previous custody < 21	No significance	No significance
Previous custody > 21	No significance	No significance
Previous custody total	No significance	No significance
No. of offence categories	No significance	No significance
No. previous violent	No significance	No significance
No. previous sexual	No significance	No significance
OASys total score	Completers < Non-starters Non-completers < Non-starters	Starters < Non-starters
OASys criminogenic score	Completers < Non-starters	Starters < Non-starters
OASys non-criminogenic score	No significance	No significance

Table 5.44: Summary of significant findings cont.

Variable	Three group (completer, non-completer, non-starter	Two group (starter, non-starter)
OASys individual needs scores	No significance	Criminal history: Starters < Non-starters
OASys individual criminogenic need scores	No significance	No significance
OASys individual non-criminogenic need scores	No significance	No significance
OASys risk classification	Fewer low risk Non-starters More high risk Non-starters	Fewer low risk Non-starters More high risk Non-starters

Discussion

The aim of this chapter was to determine whether there were any associations between a host of offence- and need-related variables and the completion status of individuals sentenced to attend the ETS programme within community settings. A number of such associations were observed amongst the data. This section will discuss these findings and their implication for future practice in relation to offending behaviour programmes.

Perhaps the first point to make is that the analyses returned few differences, either in relation to the criminal history or need variables, between the non-completers and other two groups. Indeed, the initial three group analysis only produced significant differences between the non-completers and the completers in relation to their experience of previous breach (the non-completers had more experience of breach) and between the non-completers and the non-starters in relation to their total OASys risk of reconviction score (the non-starters had higher scores[19]). Despite this, the regression analyses indicated that the completers and non-completers differed in respect of their Copas rate. An examination of the Copas rate means reveals that the non-completers' rate of offending is similar, but not quite as high, as that of the non-starters. The non-completer group, therefore, are younger than both completers and non-starters (as seen in chapter four), more likely than the completers to be aggressive (see aggressiveness scores on p. 126) or violent (see the proportion of violent index offences on p. 143 and the number of previous violent offences on p. 150), have more convictions under the age of 18 than the completers, but similar numbers of convictions across the lifespan. This latter finding is perhaps mediated by the differences in Copas rate and age: non-completers are, on average, four and a half years younger than the completers but their rate of offending is over seven points higher (and in fact is closer to that of the non-starters than that of the completers). Hence if the non-completers were

[19] This finding reveals differences in the calculation of risk or reconviction between the two tools utilised within this thesis. In chapter four, there were no differences between the OGRS2 risk of reconviction scores of the non-completers and the non-starters but here differences based on the OASys risk of reconviction scale were observed. This contradiction likely reflects the use of different predictor variables: the OGRS2 tool relies solely on static, criminal history variables whilst the OASys tool uses criminal history *and* dynamic need variables to calculate risk scores. This would imply that the criminal histories of the non-completers and non-starters are similar but that they differ in relation to their dynamic needs.

to continue at their current rate of offending, by the time they reached the age of the non-starters, it is likely that they would have similar numbers of convictions as this group. Such a finding reveals the need for Probation Areas to engage the younger, slightly more violent, and 'quick-paced' offenders within the interventions to which they initially commence to ensure that they proceed to completion and hence reduce their likelihood of reconviction.

The more pronounced differences within these analyses, however, were those between the programme completers and non-starters. Given this, it was decided to rerun the analyses to determine whether it was possible to establish differences between those who fail to commence a programme and those who at least attend session one. The distribution of previous breaches amongst these groups was found to vary. Just over half of the starters (56.7%) compared to almost three-quarters of the non-starters (70.5%) had previously been breached. As such, programme non-starters are significantly more likely to have failed to comply with a previous (or indeed current) order than the programme starters. Further consideration of the figures shows that the majority of offenders (62.5%) had previously been subject to a previous breach of probation, parole, license, bail, or a community based sentence. Such a finding demonstrates the difficulty that probation officers face in keeping offenders compliant with the order to which they have been sentenced. In addition, those who had previously breached an order were more likely than those who had not to fail to complete their programme (62.3% and 48.2% respectively). This is not to say, however, that those with a history of non-compliance should not be considered for allocation to a programme; over a third of those who had failed to comply in the past still managed to complete the programme to which they were sentenced. Instead these findings highlight the need for further, more detailed investigation; for example, is there an association between the type of the previous breach and the subsequent programme attendance? Or, does the timing of the previous breach tell us anything about whether an offender would fail to complete a programme?

In relation to the previous convictions of programme referrals, non-starters had significantly more previous court appearances at which convicted than the programme starters. Most of the differences seen, however, related to convictions received as a juvenile: there were no differences between the two groups in the number of

convictions received after the age of 18. On comparing the ages of the groups and the ages at which the groups were first convicted, this finding becomes more clear. The average age at which starters were first convicted was approximately 18 whilst the non-starters were first convicted, on average, at age 16. Programme non-starters were therefore found to commence their criminal careers at an earlier point in their lives than programme completers. In relation to offences committed after the age of 18, however, the non-starters, being younger in age than the starters (see p. 122 in chapter four), have had less time to accumulate subsequent offences. Their rate of offending, however, is significantly higher than that of those who commenced their programme. The findings here (and those above in relation to programme non-completers) would suggest, therefore, that rather than programme completers simply being more advanced on their developmental trajectories (as might be assumed by their age), the completer and dropout groups are actually distinct typologies of offenders. Programme non-starters are younger, commenced offending at an earlier age, committed more offences as a juvenile, show a greater propensity to breach their orders, have a higher rate of offending, and are more likely to commit further offences than the programme completers.

Despite these findings, there were no differences between the groups in relation to the number of previous custodial sentences they had served, either under the age of 21, above the age of 21, or in total. Both groups had been given approximately two custodial sentences over their lifespan with one occurring before the age of 21 and a second since turning 21. In addition, there were no differences between the groups in relation to the number of different types of offences committed over their lifespan or the number of violent or sexual convictions.

To summarise the findings relating to offence-related variables, the groups were found to differ in relation to some of the variables under consideration but certainly not all of them. Non-completers were younger, and showed a tendency to be more aggressive than the other two groups. However, their rate of offending was close to that of the non-starters, which was significantly higher than that of programme completers. There was an association between previous breach and dropout from programmes. Moreover, programme non-starters were likely to have more convictions as a juvenile than programme starters and started offending at a younger age than programme

completers. There are indications therefore that programme non-starters tend to lead a lifestyle that is more criminogenic and anti-authority than programme completers and that the non-completers are well on their way to the same. Despite this, the groups did not differ on all of the offence-related variables: no differences were found between the groups' experience of custody, or the type of previous offending behaviour.

It would seem therefore that there are indications that the non-starter group, and perhaps also the non-completer group, bear a resemblance to the "chronic" offenders described by Wolfgang, Figilio, andk Sellin (1972) and Farrington and West (1993) or the life-course persistent offenders illustrated by Moffit (1993). These offenders are small in number but account for a large proportion of crime, commence their offending and other forms of anti-social behaviour at an early age, and continue their offending behaviour into adulthood and beyond.

The implications of these findings for the 'would do well anyway' argument (Debidin & Lovbakke, 2005) are somewhat mixed: some of the variables associated with increased risk of reconviction were also found to be associated with attrition. However, these did not significantly predict attrition within the logistic regression, nor did all of the variables related to reconviction differentiate the groups. It would seem from this analysis therefore, that the 'would do well anyway' argument is not sufficient to explain the 'selection' of programme referrals into completers and dropouts of programmes. Having said this, caution should be expressed here: in calculating the OGRS2 score from the offence-related variables utilised in the above analyses, the variables are weighted according to the strength of the association with future reconviction (Taylor, 1999). As such, some of these variables are more predictive of reconviction than others. It is feasible, therefore, that the analyses presented above lack the power to detect differences between the groups on the more subtle predictors of reconviction.

Analysis of the OGRS2 equation (see Appendix B) reveals that the relative contributions of the variables to the overall OGRS2 score are difficult to determine in a meaningful way. This is due, in the main, to the banding of variables such as age at conviction and age at first conviction which are then weighted according to their contribution to risk of reconviction. The contributions of these variables to the equation thus depend on the banding within which each individual falls and hence do not contribute in a linear manner to the calculation of risk. For example, for those offenders

who are aged 21 and over and/or commenced offending at age 21 or over, the age information results in a protective contribution to the model and hence reduces the overall risk score. This reflects the conclusions of longitudinal researchers such as Moffitt (1993) and Farrington (1992; 2003) that most people desist from offending behaviour as they reach and progress through their twenties.

Likewise the contribution of the number of previous convictions to the model is less than straightforward. Rather than the continuous number of convictions receiving a weighting which is then added to the model, this information is only included in the model as part of the calculation of the Copas rate. Indeed, the age at first conviction data is only used to produce an indication of the length of the individual criminal career thus far. This is then used alongside the convictions data to produce the Copas rate. For the calculation of the OGRS2 risk of reconviction score then, it is the rate of conviction that contributes to the model rather than the age at first conviction or the number of convictions an individual has accumulated. As such, it is not possible to determine those variables that contribute most to the OGRS2 model as this will vary on an individual basis due to the construction of the equation. Consequently, is has not been possible to determine whether those variables that contribute most to the risk of reconviction calculation are also those that differ between the completers, non-completers and non-starters of programmes. The findings reported here should therefore only be taken as tentative evidence against the 'would do well anyway' argument.

Given the finding in the chapter three that the groups did not differ in relation to the number of criminogenic, non-criminogenic or total number of needs, it was surprising to find otherwise within these data. Both the total OASys and the criminogenic need score were found to differentiate between the groups. Analysis of the total OASys score revealed that the non-starters scored, on average, ten points higher than both the completers and non-completer groups. The criminogenic needs scores were also significant higher for the non-starters compared to the completers. There were no differences in relation to the non-criminogenic need scores implying that the difference in overall OASys score is accounted for by the differences in the criminogenic needs score only.

To find that the non-starters and non-completers differed in relation to their total OASys scores was surprising. The total OASys score is an additional risk of

reconviction assessment (Howard, 2006) and hence that it differentiated these groups when the OGRS2 risk score failed to do so was unexpected. The reason for such a finding, however, may lie in the utilisation of different variables by the two assessment tools. The OGRS2 tool uses static criminal history variables only whereas the OASys tool supplements these data with information relating to offender need. Thus it could be reasoned that, if the OGRS2 fails to detect any differences between the groups, the difference between the non-completers and non-starters in relation to the total OASys score arises due to offender need, rather than criminal history, differences between the groups. Alternatively, it could be (and this would also perhaps explain why differences between the non-completers and non-starters have been found in the criminal history variables but not in their overall OGRS2 score) that the manner in which the OGRS2 and OASys tools process the static criminal history information in their calculations differ. Examination of the groups' need scores will perhaps go some way to untangling this issue.

Analysis of the need domains, however, revealed no significant differences between the groups on any of the domain scores. Despite not reaching statistical significance[20], however, the means and effect sizes show small differences in relation to accommodation, financial management and income, education, training and employability, lifestyle and associates, drug misuse, and attitudes with the non-starters scoring, in all cases, higher than both the completers and the non-completers. As such this group do not seem to present issues in relation to one need area only but instead seem to live generally chaotic lifestyles. This finding therefore concurs with the anecdotal evidence provided within probation officer reports (Hollin, et al., 2002a, 2002b).

A note of caution should be expressed when comparing the findings relating to offender need with those reported within chapter three. The data within the previous chapter drew upon the 'mini-OASys' tool, the precursor to the full OASys tool. As such, there are differences in the calculation of the scores which do not make them comparable with those in this chapter. In chapter three, the mini-OASys tool raw scores were standardised to ensure comparability across the eleven different need scores. The

[20] This is perhaps due to a conservative use of the Bonferroni adjustment to the familywise error rate.

total score was a simple sum of the eleven standardised needs scores and the criminogenic total score was the sum of those needs which were thought, for that individual, to be related to their offending behaviour.

The OASys scores presented within this chapter, however, were calculated in quite a different manner. The OASys system requests data relating to ten needs (listed in table 5.13) and the offenders' criminal history. The system then calculates a score for each of the categories. These scores are, however, not standardised and hence comparisons across the categories are not possible. Moreover, the OASys system weights each of the need scores in accordance with the findings of large scale reconviction research: those that are found to be more predictive of reconviction are weighted more highly than those which were less predictive of reconviction. These weighted scores are finally summed to create the overall OASys total score.

The difference between the calculation of the scores within this and the previous chapter, therefore, lies in the weighting that takes place for the OASys data but not for the mini-OASys data. The aim of this weighting is to provide the service provider with an indication of the likelihood of the offender's risk of reoffending similar to that obtained with OGRS2 but calculated using static and dynamic factors. What would be more useful in relation to the body of research interested in attrition from programmes would be for future research to use standardised scores (as opposed to weighted scores) to determine which needs are of interest in relation to attrition.

The main finding to come out of these analyses therefore is that programme non-starters seem to be more criminogenic in nature and have more varied and complex needs than the programme starters within the sample. Relative to starters, non-starters had received significantly more convictions, both under the age of 18 and in total, had received their first conviction at a younger age, had been subject to more previous breach proceedings, and had higher OASys total and criminogenic scores. Programme non-completers were also more criminogenic than the programme completers but, perhaps due to their younger age, were not so far developed in this regard as the non-starters. The rate of conviction of this group, however, was similar to that of the non-starters. It is possible, therefore, that the programme non-completers are younger versions of the programme starters: as such they are slightly less criminogenic and have fewer needs but show the signs of heading in a similar direction

to the non-starters if they do not desist from offending. Indeed, the findings from longitudinal studies of crime indicate that “people who commit relatively many offenses during one age range have a high probability of also committing relatively many offenses during another age range” (Farrington, 2003, p. 223). It is therefore crucial that the Probation Service ensure that they do all they can to motivate and support these offenders through the programmes to which they are referred.

Chapter Six

An investigation of process related variables and their relationship with attrition

Introduction

The previous two chapters have examined the relationships between attrition from offending behaviour programmes and individual factors such as age, risk of reconviction, and criminal history. At times within these chapters, the way that the probation area used these data in the procedures relating to offending behaviour programmes referrals has been assessed. However, there is another set of variables which have the potential to impact on programme completion and dropout which are based entirely within the control of either the judiciary or the correctional service. These factors relate to the processing of the offender through their sentences and their offending behaviour programmes. Utilising the same dataset as the previous two chapters, this chapter will investigate the relationships between process or organisational factors and programme completion and dropout. The factors under examination include the type of order that the offender is subject to, the length of time from conviction to the start of the programme, and the sequencing of the programme.

In their discussion of treatment engagement, Howells and Day (2007) state that "The most dramatic and, arguably, significant form of low engagement is failure to complete the programme" (p. 48). This may be true; failing to complete a programme may indeed signify that the offender has not engaged with the treatment provision. However, caution should be expressed when evaluating such a statement. Such assertions could be taken as support for the view that all attrition is related to, or indeed caused, by low engagement. Such a position, however, perhaps ignores the influence of situational factors on attrition: if an offender is ill or has an appointment at the job centre which he or she chooses to prioritise over attendance at an allocated intervention, are these signs of low engagement in the programme? Not necessarily. However, perhaps more concerning is the possibility that Howell and Day's statement could be interpreted as placing the sole responsibility for attrition with the offender. As such responsivity, organisational and programme factors are absolved of any responsibility.

It is important to explain that Howells and Day (2007) does not ascribe to the view that all responsibility for attrition sits within the offender. Instead they attest to the notion of 'treatment readiness' which they define below:

> The concept of readiness can be broadly defined as the presence of characteristics (states or dispositions) within either the client or the therapeutic situation, which are likely to promote engagement in therapy and which, thereby, are likely to enhance therapeutic change. (p. 49)

As such the authors situate responsibility for low engagement, and hence attrition, within the individual offender and within the therapeutic situation. This chapter will attempt to determine which of a number of organisational factors are associated with attrition from programmes. It will also test the above statement to determine whether low engagement with, and understanding of, the programme is indeed associated with within-treatment attrition.

The first factor that will be considered within this chapter in relation to attrition is the order type that the offender is subject to. During the period to which the data refers, sentenced offenders (as opposed to those who may be participating in programmes whilst on licence from prison) could either be subject to a community order[21] or a suspended sentence order, with the added requirement of attending an offending behaviour programme. Both orders are administered by the Courts but they differ in that the suspended sentence order should only be imposed when the 'custody threshold' is passed. This threshold is passed when:

> The Court …is of the opinion that the offence, or the combination of the offence and one or more of the offences associated with it, was so serious that neither a fine alone nor a community sentence can be justified for the offence. (Criminal Justice Act, 2003, Section 152 (2))

[21] Under the Criminal Justice Act 2003 (and from April 2005) probation orders have the generic term of 'Community Orders'. An accredited programme can be undertaken as a requirement of the community order. Prior to this date, however, offenders on programmes were either on a community rehabilitation order or a community punishment and rehabilitation order (or indeed a suspended sentence). As the data collection period spans this change, to avoid confusion the generic 'community order' terminology will be used throughout this chapter.

As such, offenders subject to a suspended sentence order are provided with a term which they must serve in the community and a prison sentence term. Should the community part of the order be breached, the prison sentence can be invoked and hence the offender will serve their prison term.

Given that the consequence for breaching a suspended sentence order is potentially more severe than that of breaching a community order, it is possible that offenders on a suspended sentence order feel more coerced to attend their programme as a result of this additional pressure. Despite prison research finding no relationship between perceptions of coercion and treatment retention (Kolton, 2004), Young (2002) has reported that research into the effects of legally mandated interventions on treatment retention and outcome has been supportive of the use of coercion to reduce attrition: "Coercion might... enhance retention in treatment programs that have long had problems with high drop-out rates" (Young, 2002, p.28). However, in his own research on a community-based drugs intervention, Young reported that different types of coercion can produce different outcomes: perceived legal pressure was found to predict treatment retention and, more specifically, explaining the requirements of the treatment, the consequences of dropout and convincing the offender that these consequences will be enforced were effective in terms of improving retention in treatment. However, no support was found for threatening or imposing prison sentences on the programme participants in an attempt to increase retention.

Given these findings it is unclear if a suspended sentence order would be expected to increase retention rates over and above those of offenders on a community order. Even though the consequences of breaching the suspended sentence order are explained to the offender by his or her offender manager at the beginning of a sentence, the findings of Young (2002) imply that the threat of prison might outweigh this knowledge. This chapter will therefore investigate whether there is an association between order type and attrition from the ETS programme.

In chapter three, the targeting of offenders for general offending behaviour programmes according to their OGRS2 risk of reconviction scores was investigated. The findings within chapter three were consistent with those of Palmer et al (2008) and Palmer et al (2009): despite a clear criterion for selection of offenders for programmes

based on their OGRS2 scores, a large proportion of offenders did not meet this criterion. In chapter three, 43.4% of offenders who were referred to one of the two general offending behaviour programmes had OGRS2 scores which fell outside of the specified range (9.2% fell below and 34.1% fell above the specified range). These figures are similar to those reported by Palmer et al (2008) and Palmer et al (2009). Given that Palmer et al (2008) also reported that the interaction between retention in treatment and appropriateness predicted reconviction outcomes with non-completers in the too low group performing worse than expected and completers within the too high group performing better than expected, this chapter will repeat these analyses to determine whether the previously observed patterns relating to OGRS2 appropriateness have altered in the time between the two data collection phases (2002 in chapter three to 2005-6 in this chapter).

One of the Correctional Service's Accreditation Panel criteria for the accreditation of programmes states that interventions for offenders should specify their sequencing, intensity and duration (Home Office, 2003b) and these should be tailored to the risk level of the offenders targeted for the programme. Whilst the risk principle (Andrews & Bonta, 2006) specifies that higher risk individuals should be allocated to programmes of longer duration, there is little research that has assessed the impact of the intensity of the delivery of offending behaviour programmes on attrition or treatment outcomes. Drawing on research relating to smoking cessation programmes, high intensity interventions have been shown to have more positive immediate effects; however, no differences in outcomes between low and high intensity programmes remained after six months (Fiore, Bailey, Cohen, Dorfman, Goldstein et al., 1996; Silagy, Mant, Fowler, & Lancaster, 1999). However, no information relating to the relationship between intensity and attrition from these programmes was provided. This chapter will therefore assess whether there is any association between the intensity or pace of delivery of the ETS programme and attrition. The Probation Area from which the data were collected varies the pace of the delivery of the ETS programme depending on timetabling arrangements and demand for the programmes. As such, the ETS programme was delivered across the data collection period either twice or thrice weekly. This chapter will compare the attrition rates between the two delivery paces to assess for any patterns in the data.

The National Offender Management Service is governed by a set of Offender Management National Standards which "cover the underlying management process which is applicable to all offenders" (Ministry of Justice, 2007, p. 4). Within these Standards, and also within the audit criteria associated with offending behaviour programmes (National Offender Management Service, 2007), it is specified that general offending behaviour programmes should commence[22] no later than six weeks after sentence or release on licence. However, delays in commencement on the programme can be deemed acceptable if work is being undertaken to increase motivation or resolve immediate issues within the offender's lives. Very little research to date has assessed the effect of length of time between receiving the court order and commencement on the programme on attrition from programmes. Following an exhaustive review of the research, only two studies were found to assess this directly. Turner (2006) studied attrition from four accredited programmes delivered within community settings within England. He found that programme completers waited a significantly shorter period of time to comment their programme than non-completers. However, despite National Standards the programme completers still waited, on average, 121 days prior to starting their programme. However, in contrast Chu's (2002) assessment of attrition from the Think First programme, again within community settings, found no association between time from order to commencement and attrition. This chapter will therefore evaluate whether there are any differences between the 'time to programme start' periods of programme completers, non-completers and non-starters.

Finally, this chapter will also assess whether the offenders' levels of engagement with, and their levels of understanding of, the ETS programme are associated with their likelihood of programme completion. To return to the comments of Howells and Day (2007), they propose that programme attrition is the ultimate demonstration of low engagement with the intervention. Whilst not strictly relating to process factors (although of course it is feasible that low engagement is in part a product of process factors), this hypothesis will be tested within this chapter due to the relevance it has to organisational practices. Within this context, the engagement and understanding ratings are provided by the programme tutors following each session. If, therefore, the tutors'

[22] A commencement is defined as attendance at session one of the core programme.

ratings of offender engagement, or indeed understanding, are found to be associated with attrition then this has implications for the organisation. If an association between attrition is found and tutors witness low engagement or understanding within their sessions, extra motivational work or support could be provided to those individuals most at risk of dropout. Of course, this data analysis relates solely to in-programme attrition as ratings of engagement and understanding are collected after each programme session and hence are not available for those who do not commence the programme. However, given the findings within chapter five relating to the programme non-completers' criminal history and rate of offending, the likelihood is that without intervention these offenders will continue on with their criminal careers. As such, if tutor ratings can be used to identify those offenders most at risk of becoming non-completers, adapting the process in such a way as to support their ongoing attendance on the programme may be possible.

This research will therefore evaluate whether there is an organisational or process element to attrition. Given the conclusions of the early process evaluations of offending behaviour programmes within England and Wales that the observed high levels of attrition were associated with implementation failure (Hollin et al., 2004), it is not unreasonable to assume that some of the factors considered here will be associated with attrition from, or completion of, the ETS programme. The findings will be discussed with reference to their practical and theoretical implications.

Hypotheses:

1. There will be a significant association between order type (community order, suspended sentence order) and group (completer, non-completer, non-starter) with those on suspended sentence orders being more likely than those on a community order to complete their programme.
2. There will be a significant association between OGRS2 appropriateness (too low, appropriate, too high) and group (completer, non-completer, non-starter): there will be a higher proportion of 'too low' offenders within the completer group comparative to the dropout groups and a higher proportion of 'too high' offenders within the dropout groups compared with the completer group.

3. There will be a significant difference between the groups' (completer, non-completer, non-starter) time to programme start. The dropout groups will have waited longer between receiving their order and starting the programme than the completers.
4. There will be a significant difference between the groups' (completer, non-completer, non-starter) level of engagement scores. The completers will have significantly higher level of engagement scores than the dropout groups.
5. There will be a significant difference between the groups' (completer, non-completer, non-starter) level of understanding scores. The completers will have significantly higher level of understanding scores than the dropout groups.

The data will also be explored to test for any associations between the number of attempts at programme start on the order, the time of day that the ETS programme is delivered, and group (completer, non-completer, non-starter).

Method

Design

This study utilised the same dataset as the previous two chapters. This time the aim was to determine whether the groups (completers, non-completers, non-starters) differed in respect to process related variables. The independent variable within these analyses, hence, was the group which again comprised of three levels: completer, non-completer, non-starter. The dependent variables this time all related to organisational or process variables.

Participants

As in the previous two chapters the participants were 293 offenders from one Probation Area sentenced between 1st January 2005 and the 31st December 2006 to a community sentence with the addition requirement of attendance at the ETS offending behaviour programme. The participant descriptives were therefore the same as the previous two chapters: the average age of those in the study was 25.56 (*SD* = 7.89) with a range of ages from 18 to 55 years; the mean OGRS2 score was 64.42 (*SD* = 21.29) with

a range scores from 10 to 99; and the mean number of previous convictions was 8.27 (*SD* = 7.046) with a range from 0 to 44. Of the sample, 27 (9.2%) were female and 266 (90.8%) were male. In relation to the offenders' type of sentence, 209 (71.3%) had been sentenced to a community order and 84 (28.7%) a suspended sentence order. Of the sample, 122 (41.6%) failed to commence the ETS programme (non-starters), 45 (15.4%) commenced the programme but failed to complete it (non-completers), and 126 (43%) completed the programme (completers).

Measures

As explained within previous chapters, data were obtained from a variety of Probation electronic resources: CRAMS, OASys, and IAPS. In relation to the creation of the independent variable, the IAPS system allowed classification of the offenders into programme completers, non-completers, and non-starters. The majority of the dependent variable data was obtained from the IAPS database which holds information relating to the delivery of programmes. The remaining data, such as the offenders' order type and their date of sentence (used to calculate the time from order to the commencement of the programme) was obtained from the CRAMS database.

Procedure

The procedure used in the collection of the data for analysis within this chapter was similar to as described in detail in chapter four. The majority of the data were collected either from the CRAMS database (order type, OGRS2) or IAPS (number of attempts on this order, number of programme sessions a week, evening/day programme, level of engagement and understanding). The OGRS2 appropriateness variable was calculated by comparing the OGRS2 risk of reconviction score to the targeting criteria for the general offending behaviour programmes. Scores of 30 and below were classified as 'too low', from 31 to 74 as 'appropriate' and 75 and above as 'too high'. Finally, the 'time from order to programme start' variable was calculated by measuring the time, in days, from the date of conviction to the start date of the first programme to which the offender was referred.

Results

Order Type: Table 6.1 displays the distribution of order types amongst the three groups. As can be seen, a greater proportion of those undertaking the ETS programme whilst on a suspended sentence completed the programme compared to those on a community order. A greater proportion of completers were on suspended sentences compared to the drop-out groups. This might indicate that those on suspended sentences, having more to lose if they are breached for non-attendance or feeling more coerced, tend to be more compliant with the order. This distribution, however, was not statistically significant: χ^2 (2, N = 293) = 3.787, p = 0.151, Φ_c = 0.11.

Table 6.1: Order type by group

	Community Order (%)	Suspended Sentence Order (%)	Total (%)
Completers	83 (39.71)	43 (51.20)	126 (43.00)
Non-completers	32 (15.31)	13 (15.48)	45 (15.36)
Non-starters	94 (44.98)	28 (33.33)	122 (41.64)
Total	209 (100.00)	84 (100.00)	293 (100.00)

Number of attempts on this order: Some participants were permitted more than one attempt on the ETS programme. Table 6.2 presents the mean number of attempts at ETS on their present order by group. In this context, an attempt is counted when the participant is allocated to a programme. As such, a large proportion (N=76, 62.29%) of non-starters were designated as having no attempts at a programme as they did not progress as far as the allocation stage hence the value possible on this variable ranged from 0 upwards.

These data were non-normally distributed (completers: *K-S Z* = 4.468, $p < 0.001$, non-completers: *K-S Z* = 1.815, $p = 0.003$, non-starters: *K-S Z* = 3.972, $p < 0.001$) hence a non-parametric Kruskall-Wallis was conducted on the data to test for differences between the groups. This test indicated significant differences between the groups and

a strong association between group and the number of attempts: χ^2 (2, N = 293) = 113.52, $p < 0.001$, $\Phi_c = 0.62$. Post-hoc Mann-Whitney tests revealed that the Non-starters had significantly fewer attempts than the Completers (U = 2663, $p < 0.001$) and the Non-completers (U = 748, $p < 0.001$), and that the Completers had fewer attempts than the Non-completers (U = 2303, p = 0.031).

Table 6.2: Mean number of attempts by group

	Attempts
Completers	1.49[a] (0.80[b])
Non-completers	1.76 (0.88)
Non-starters	0.52 (0.78)
Total	1.13 (0.96)

[a] Mean [b] Standard Deviation

The maximum number of attempts within the sample was four. Table 6.3 presents the cumulative completion status outcomes in relation to each attempt. As expected, as the number of attempts increase, so do the number of completers whilst the number of non-starters and non-completers decrease. The final column in this table presents the gain in completers per attempt, first as a number and second as a percentage of the total number of attempts at this stage. A note of caution should be expressed here, however; this relationship is not as straightforward as it might look. It is not the case, for example, that all 14 individuals who moved out of the non-starters group at the second attempt became completers; some may have become non-completers and hence more than nine non-completers may have become completers. Despite this shortcoming, the table presents the overall completion status figures so the overall gain at each attempt can be established.

Table 6.3: Cumulative completion status outcomes in relation to the number of attempts.

	Cumulative group numbers				
	N	Completers	Non-completers	Non-starters	Completer gain (% of N)
First attempt	293	85 (29.0%)	65 (22.2%)	143 (48.8%)	-
Second attempt	77	110 (37.5%)	54 (18.4%)	129 (44.0%)	25 (32.47)
Third attempts	27	122 (41.6%)	48 (16.4%)	123 (42.0%)	12 (44.44)
Fourth attempt	8	126 (43.0%)	45 (15.4%)	122 (41.6%)	4 (50.00)

OGRS2 appropriateness: As mentioned previously, the targeting criteria for the ETS programme states that offenders should have an OGRS2 score between 31 and 74. Analyses were undertaken to assess whether these criteria had been adhered to and whether there were any patterns in the data in relation to this. Each offender was classified based on their OGRS2 score into two groups: those who fell between 31 and 74 ('appropriate) and those who did not ('not appropriate'). Table 6.4 shows the distribution of this classification between the three groups. With totals very similar to those discussed within previous chapters, almost one in two offenders fell outside of the OGRS2 targeting criteria (43.3%). Exactly the same proportion of non-starters and non-completers were deemed to be appropriately selected based on their OGRS2 scores. A greater proportion of completers fell within the appropriate range, however. The chi-square analysis found this distribution to be non-significant, χ^2 (2, N = 293) = 1.787, p = 0.409, Φ_c = 0.08.

Table 6.4: OGRS2 criteria classification by group

	Appropriate (%)	Not appropriate (%)	Total (%)
Completers	77 (61.1)	49 (38.9)	126 (100)
Non-completers	24 (53.3)	21 (46.7)	45 (100)
Non-starters	65 (53.3)	57 (46.7)	122 (100)
Total	166 (56.7)	127 (43.3)	293 (100)

To repeat the analysis within chapter three with these data, a further analysis classified the offenders into three groups according to their OGRS2 scores and the targeting criteria for the ETS programme: 'too low' (30 and below), 'appropriate' (31 to 74), and 'too high' (75 and above). Table 6.5 presents the distribution across groups. As can be seen, this distribution is very similar to that within chapter three. Relative to the non-completer and non-starter groups, the completer group comprises a much larger proportion of 'too low' offenders (2.5% non-starters, 4.8% non-completers and 13.5% completers). Additionally, the drop-out groups have a greater proportion of 'too high' offenders comparative to the completers (44.4% and 44.4% compared with 25.4%). This distribution was significant, χ^2 (4, N = 173) 20.218, $p < 0.001$, $\Phi_c = 0.19$.

Table 6.5: OGRS2 appropriateness by group

	Too low (%)	Appropriate (%)	Too high (%)	Total (%)
Completers	17 (80.95)	77 (46.39)	32 (30.19)	126 (43.00)
Non-completers	1 (4.76)	24 (14.46)	20 (18.87)	45 (15.36)
Non-starters	3 (14.29)	65 (39.16)	54 (50.94)	122 (41.64)
Total	21 (100.00)	166 (100.00)	106 (100.00)	293 (100)

Time from Order to First Programme Start: In order to determine any relationship between the time from order to programme start and group, it was

necessary to use only those data which related to individuals allocated to a programme. As mentioned above, a large proportion of non-starters did not reach the allocation stage and hence were not given a programme start date. As such these were excluded from the analysis. In addition, programme start date information was missing for one completer hence this individual was also excluded from the analysis. The final dataset for this analysis therefore comprised 216 participants: 85 completers, 65 non-completers, and 66 non-starters.

As a proportion of participants had more than one attempt at the ETS programme, it was necessary, for the purpose of this analysis, to compare the time from order to their first allocated programme. In addition, the completion status used within this analysis relates to the participants' first attempt at the ETS programme.

From the order and programme start dates, the number of days from the offenders receiving the order to the start of their first programme was calculated. Table 6.6 presents the mean number of days from the date of the Probation Order to the start of the allocated programme by group.

Table 6.6: Mean time from order to programme start by group.

	Time from order to programme start in days
Completers (N = 85)	90.58[a] (62.01[b])
Non-completers (N = 65)	91.21 (66.99)
Non-starters (N = 66)	81.17 (43.35)
Total	87.89 (58.54)

[a] Mean [b] Standard Deviation

These data were non-normally distributed with a negative skew (completers: *K-S Z* = 1.670, *p* = 0.008, non-completers: *K-S Z* = 1.561, *p* = 0.015, non-starters: *K-S Z* = 1.437, *p* = 0.032). Such distributions are to be expected given the practical issue to which the data refer: most offenders would be allocated to a programme soon after sentencing with a small tail of those having to wait longer for operational or personal reasons. As

the normality assumption of parametric tests had been violated, a Kruskall-Wallis test was conducted to test for differences between the groups. No significant differences were observed: χ^2 (2, *N* = 216) = 0.214, *p* = 0.899, Φ_c = 0.03.

Time of first programme (day or evening): Within the probation area studied, the ETS programme is delivered as a day and an evening programme. Using the same dataset as described above in relation to the time from order to programme start analysis, there was an exact split between the numbers allocated to a day or an evening programme. Table 6.7 presents the distribution of group by time of first programme. The data demonstrates a slight trend towards higher completion and lower non-completion of the day programme relative to the evening programme. However, the day programme also suffered from a greater proportion of non-starters. This distribution, however, was not statistically significant: χ^2 (2, *N* = 216) = 1.783, *p* = 0.410, Φ_c = 0.09.

Table 6.7: Distribution of group by time of programme.

	Day programme (%)	Evening programme (%)	Total (%)
Completers	45 (41.67)	40 (37.04)	85 (39.35)
Non-completers	28 (25.93)	37 (34.26)	65 (30.09)
Non-starters	35 (32.41)	31 (28.71)	66 (30.55)
Total	108 (100.00)	108 (100.00)	216 (100.00)

Number of programme sessions per week: The delivery of the programme can also vary within the probation area studied by the number of sessions delivered per week. Programmes are delivered either in a two or a three sessions per week format. The distribution of the group numbers across the two modes of delivery are displayed in table 6.8. Using the dataset relating to the first programme allocation as described above once more, 178 (82.4%) were allocated to a programme which was delivered twice weekly whilst the remaining 37 (17.1%) were allocated to a thrice weekly programme. Data were missing for one offender.

Table 6.8: Distribution of group numbers of programme sessions per week.

	Two sessions (%)	Three sessions (%)	Total (%)
Completers	67 (37.64)	18 (48.65)	85 (39.53)
Non-completers	57 (32.03)	8 (21.62)	65 (30.23)
Non-starters	54 (30.34)	11 (29.73)	65 (20.23)
Total	178 (100.00)	37 (100.00)	215 (100.00)

The percentage of non-starters remained almost static across the two rates of delivery. However, the percentages of completers was higher and non-completers was lower when the programme was delivered thrice weekly compared to when delivered twice a week. This distribution, however, was not statistically significant: χ^2 (2, N = 215) = 2.039, p = 0.361, Φ_c = 0.10.

Level of offender engagement and understanding: Each offender's level of engagement with and understanding of the programme is rated by the programme facilitators after each session of the programme. Each construct is scored on a scale of one to five. Each offender's scores were averaged across the number of sessions they attended and are displayed in table 6.9 below. As such, these data were not available for non-starters of the programme.

Table 6.9: Mean Level of Engagement and Understanding scores by group.

	Engagement	Understanding
Completers (N = 85)	3.21[a] (0.57[b])	3.18 (0.60)
Non-completers (N = 65)	2.86 (0.71)	2.80 (0.71)
Total	3.12 (0.63)	3.08 (0.65)

[a] Mean [b] Standard Deviation

The non-completers understanding scores were non-normally distributed (*K-S Z* = 1.423, *p* = 0.035) and hence the differences between completers and non-completers were investigated using a non-parametric Mann Whitney test. The differences were significant with completers having significantly higher understanding scores than the non-completers (*U* = 1925.5, *p* = 0.006). An independent samples t-test investigated the differences between the completers and non-completers engagements score. As with the understanding scores, the completers scored significantly higher than the non-completers (*t* [166] = -3.277, *p* = 0.001).

Prediction of group from variables which significantly vary between groups:

A multinomial logistic regression was performed to determine whether it was possible to predict group membership (completer, non-completer, non-starter) from the number of attempts on this order and the OGRS2 appropriateness classification. Whilst the two variables concerned with level of engagement and level of understanding also differed between the groups, these data were not available for the non-starters. As such, the multinomial regression assessed the prediction of the three groups from the number of attempts on this order and the OGRS2 classification whilst a later logistic regression will add in the offenders' level of engagement and understanding variables.

For the multinomial logistic regression, the fit to the model on the basis of the entered predictors was not good, χ^2 (22, N = 293) = 81.227, $p < 0.001$, using a deviance criterion. Comparisons of the log-likelihood ratios for models with and without predictors, however, showed a reliable improvement with the addition of the predictors (χ^2 (6, N = 293) = 133.091, $p < 0.001$). Correct classification rates were 84.1% for the Completers, 4.4% for the Non-completers condition, and 73.0% for the Non-starters condition. The overall correct classification was 67.2%, which represents an increase of 33.87% over by-chance accuracy. Analysis of the contribution of individual predictors to the model with and without each predictor revealed that both predictors significantly predicted the outcome: number of attempts on this order, (χ^2 (2, N = 293) = 112.202, $p < 0.001$) and OGRS2 classification, (χ^2 (4, N = 293) = 15.454, $p < 0.001$). Analysis of the odds ratios (see table 6.10) indicates that with each unit increase in the number of attempts on this order the odds of being in the non-completer group compared to the non-starter group increase nine-fold ($p < 0.001$) whilst the odds of being in the non-

starters groups compared to the completers groups decrease by 86% ($p < 0.001$). Further analysis reveal that compared to the too high group, the too low group were less likely to be non-completers than completers (OR = 0.094, $p = 0.027$) and were less likely to be non-starters than completers (OR = 0.141, $p = 0.009$). Additionally, compared to the too high group, the appropriate group were less likely to be non-starters than completers (OR = 0.503, $p = 0.038$).

Table 6.10: *Multinomial logistic regression of programme completion as a function of number of previous attempts on this order and OGRS2 classification (too low, appropriate, too high).*

		B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
Non-	Intercept	-0.941	0.428	4.828		
completers vs.	Number of attempts	0.277	0.184	2.272	1.320	0.920 – 1.893
Completers	OGRS2 class = 0	-2.367	1.070	4.889*	0.094	0.012 – 0.764
(ref)	OGRS2 class = 1	-0.656	0.372	3.111	0.519	0.250 – 1.076
	OGRS2 class = 2	0[a]				
Non-starters vs.	Intercept	2.369	0.371	40.836		
Completers	Number of attempts	-1.997	0.277	51.890****	0.136	0.079 – 0.234
(ref)	OGRS2 class = 0	-1.959	0.754	6.745*	0.141	0.032 – 0.618
	OGRS2 class = 1	-0.688	0.332	4.295*	0.503	0.262 – 0.963
	OGRS2 class = 2	0[a]				
Non-	Intercept	-3.310	0.466	50.563		
completers vs.	Number of attempts	2.274	0.310	53.881****	9.719	5.296 – 17.838
Non-starters	OGRS2 class = 0	-0.408	1.251	0.106	0.665	0.057 – 17.838
(ref)	OGRS2 class = 1	0.032	0.412	0.006	1.033	0.460 – 2.315
	OGRS2 class = 2	0[a]				

[a] This parameter is set to zero because it is redundant

* $p < 0.05$ ** $p < 0.01$ **** $p < 0.001$

A further regression was performed to examine the effect of OGRS2 appropriateness, the number of attempts on this order, average level of engagement and average level of understanding in the prediction of completion or non-completion of the programme. There were no data in relation to the non-starters to be entered into this analysis as they had not attended the programme and hence had not received engagement and understanding scores. As such, there were only two outcome categories (completer, non-completer) and hence logistic regression could be used. As multicollinearity between predictor variables can cause problems in multivariate analysis, the intercorrelations between the variables were examined. The magnitude of most of these correlations did not exceed the level (>0.70) that would point to multicollinearity (Tabachnick & Fidell, 2006). However, the correlation between average level of engagement and the average level of understanding scores was 0.91. Given this, of these two variables only the average level of engagement score was entered into the regression. This variable was chosen over the average level of understanding score as this variable had distinguished the groups to a larger extent within the univariate analysis.

The variables entered into the analysis produced a good model fit as measured by the Hosmer and Lemeshow Test: χ^2 (8, N = 168) = 6.601, p = 0.580. This model was significantly better than a constant-only model containing only the intercept but no predictor variables: χ^2 (2, N = 168) = 20.687, p < 0.001. Classification was good with 73.8% correct classification of cases, although classification of completers (94.4%) was better than that of non-completers (14.0%). Table 6.11 shows how the predictor variables contributed to the model, along with the Wald and Exp (B) statistics for the variables. The odds ratio indicates that as the average level of engagement is increased by one unit, the odds of being a programme completer, relative to being a programme non-completer, more than double. The OGRS2 classification variable also added significantly to the predictive model; the odds ratio indicates that the odds of an offender with a 'too low' OGRS2 score being a programme completer are 3.825 times as large as the odds an offender with a 'too high' OGRS2 score being a programme completer.

Table 6.11: Logistic regression of group (completer, non-completer) as a function of OGRS2 classification (too low, appropriate, too high), number of attempts on the order, and average level of engagement scores (completers, dropouts)

	B	SE of B	Wald	Exp (B)	95% CI for Exp (B)
OGRS2 class			6.143*		
OGRS2 class (1)	2.031	1.081	3.520	7.611	0.916 – 63.425
OGRS2 class (2)	0.765	0.390	3.825*	2.148	0.999 – 4.617
Number of attempt	-0.365	0.216	2.839	0.694	0.454 – 1.061
Av. Engagement	0.870	0.310	7.889**	2.387	1.301 – 4.382
Constant	-1.558	1.025	2.309	0.211	

* $p < 0.05$ ** $p < 0.01$

Investigation of factors associated with attrition within the OGRS2 appropriateness classifications (too low, appropriate, too high):

Given the finding of an association between OGRS2 appropriateness and programme completion, it was decided to investigate further those factors that may impact on programme completion or dropout within the different OGRS2 classification groups (too low, appropriate, too high). As such, the full dataset was split into the OGRS2 classification groups and tests were undertaken to determine whether, within these groups, there were any significant differences between the programme completer, non-completers and non-starters in relation to individual, psychometric, criminogenic, and need variables. Given the small numbers within the 'too low' category, however, (there was only one non-completer and three non-starters within the too low category), this category was not included within these analyses.

The variables tested were offender age, OGRS2 score, the pre-programme psychometrics, Copas rate, total OASys score, total OASys criminogenic score and total OASys non-criminogenic score. These were chosen as they were previously shown to discriminate the attrition groups and are routinely collected by probation staff.

Age: As Kolmogorov-Smirnov tests indicated that the parametric assumption of normality had been violated by at least one group within each appropriateness category (appropriate non-starter, *K-S Z* = 1.431, *p* = 0.033; appropriate completer, *K-S Z* = 1.452, *p* = 0.029; too high non-starter, *K-S Z* = 1.423, p = 0.035; too high non-completer, *K-S Z* = 1.499, *p* = 0.022) two non-parametric Kruskall-Wallis tests were undertaken, one with the appropriate category and one with the too high category to determine whether there was a significant difference in age between programme completers, non-completers and non-starters.

Within the too high category, no significant differences were found between the groups: χ^2 (2, *N* = 106) = 2.517, *p* = 0.284, Φ_c = 0.15. However, within the appropriate category, significant differences were observed: χ^2 (2, *N* = 166) = 9.528, *p* = 0.009, Φ_c = 0.24. Post-hoc Mann Whitney U tests revealed that there were significant differences between the non-starters and non-completers (*U* = 557, *p* = 0.039) and the non-completers and completers (*U* = 548.50, *p* = 0.003). There were no significant differences between the completers and the non-starters (*U* = 2156.50, *p* = 0.156). Analysis of the means (table 6.12) indicated that non-completers within the appropriate risk category were significantly younger than the completers and the non-starters.

Table 6.12: Age by completion and appropriateness groups

	Completers	Non-completers	Non-starters
Appropriate	27.61[a] (8.29[b]) [77[c]]	22.42 (5.68) [24]	25.55 (7.05) [65]
Too High	23.66 (5.68) [32]	22.70 (8.09) [20]	22.46 (4.87) [54]

[a] Mean [b] Standard Deviation [c] N

OGRS2 score: Two one way ANOVAs were undertaken on OGRS2 scores of the completers, non-completers and non-starters: one for the appropriate category of offenders and the other for the too high group. Kolmogorov-Smirnov tests indicated no concerns in relation to normality and the Levene's test for equality of variances indicated no concerns in relation to homogeneity of variance for either the appropriate (F = 1.163, *p* = 0.315) or the too high group (F = 1.050, *p* = 0.354). No significant differences were found between the completion groups in the appropriate group, F =

0.091, *df* = 2, 163, *p* = 0.913. The ANOVA relating to the too high group approached significance, F = 3.047, *df* = 2, 103, *p* = 0.052. The means are displayed in table 6.13 below.

Table 6.13: OGRS2 by completion and appropriateness groups

	Completers	Non-completers	Non-starters
Appropriate	55.65[a] (12.27[b]) [77[c]]	56.04 (10.11) [24]	54.97 (12.49) [65]
Too High	84.75 (5.31) [32]	86.95 (5.78) [20]	88.02 (6.33) [54]

[a] Mean [b] Standard Deviation [c] N

Pre-programme psychometrics: Kolmogorov-Smirnov tests were undertaken on the pre-programme psychometric data to check for normality. These tests indicated that the parametric assumption had been violated by at least one appropriate group for the Cognitive Indolence variable (appropriate non-completer, *K-S Z* = 1.651, *p* = 0.009) and at least one appropriate group for the Racial Intimacy variable (appropriate completer, *K-S Z* = 1.609, *p* = 0.011). As such the tests assessing for differences between the completers, non-completers and non-starters for the appropriate category only for these two variables were non-parametric Kruskall-Wallis tests. No significant differences between the groups were found in relation to the Cognitive Indolence variable. Significant differences were found, however, between the groups in relation to the Racial Intimacy variable: χ^2 (2, *N* = 129) = 13.723, *p* = 0.001, Φ_c = 0.33. Post hoc Mann Whitney U tests indicated there were significant differences between the non-starters and the completers (*U* = 597.50, *p* = 0.001) and the completers and the non-completers (*U* = 424.50, *p* = 0.019). There were no significant differences between the non-starters and the non-completers (*U* = 248.50, *p* = 0.451). Analysis of the means indicates that the programme completers within the appropriate category had higher Racial Intimacy scores than the non-starters and the non-completers.

All other variables met the assumptions of normality and homogeneity of variance (assessed using the Levene's statistics for each variable within each appropriateness group). A series of ANOVAs were therefore undertaken to test for differences between the completion groups. Within the appropriately targeted group,

the Victim Hurt Denial variable (F = 3.904, *df* = 2, 116, *p* = 0.023) and the Aggressiveness variable (F = 3.309, *df* = 2, 123, *p* = 0.040) was seen to vary significantly between groups. Post-hoc Scheffe tests indicated that programme completers had significantly lower Victim Hurt Denial scores than programme non-completers (*p* = 0.001). Whilst Scheffe tests did not reveal any significant differences between the pairwise comparisons in relation to the Aggressiveness scores, the difference between the completers and non-starters approached significance: the completers had lower scores than the non-starters. Caution should be urged in relation to these findings, however; given the number of psychometric variables tested a Bonferroni adjustment would render the results non-significant.

Within the too high group, none of the ANOVAs indicated significant differences between the completer, non-completer and non-starter groups (see table 6.14).

Table 6.14: Pre-programme psychometrics by completion and appropriateness groups

		Completers	Non-completers	Non-starters
Impulsivity	Appropriate	10.92[a] (4.94[b]) [63[c]]	13.00 (3.90) [21]	12.06 (4.46) [32]
	Too High	14.03 (4.90) [29]	11.23 (4.07) [13]	13.61 (5.11) [23]
Socialisation	Appropriate	20.63 (5.62) [64]	21.28 (4.81) [18]	20.59 (6.20) [27
	Too High	21.48 (4.24) [25]	19.25 (4.34) [16]	20.20 (3.72) [20]
Locus of Control	Appropriate	42.61 (7.12) [71]	42.55 (8.90) [22]	42.03 (6.08) [32]
	Too High	42.79 (6.28) [29]	43.42 (7.61) [17]	42.67 (6.84) [21]
General Attitudes to Offending	Appropriate	33.60 (11.69) [72]	32.00 (8.89) [22]	32.97 (12.13) [32]
	Too High	39.46 (12.06) [28]	40.69 (12.90) [16]	39.87 (15.69) [23]
Anticipation of Reoffending	Appropriate	10.77 (4.36) [73]	9.91 (4.44) [23]	10.76 (5.12) [33]
	Too High	12.21 (5.49) [29]	14.29 (5.17) [17]	13.39 (5.59) [23]
Victim Hurt Denial	Appropriate	6.15 (3.03) [75]	8.29 (3.58) [21]	6.48 (2.98) [33]
	Too High	6.33 (3.02) [30]	5.18 (2.72) [17]	6.23 (3.39) [22]
Multi-culturalism	Appropriate	19.94 (4.19) [71]	19.60 (5.38) [20]	18.36 (3.60) [28]
	Too High	18.69 (5.01) [29]	19.93 (2.27) [14]	17.26 (4.08) [23]

[a] Mean [b] Standard Deviation [c] N

Table 6.14: Pre-programme psychometrics by completion and appropriateness groups (cont.)

Racial Intimacy	Appropriate	19.22 [a] (4.27 [b]) [69 [c]]	16.68 (3.41) [19]	15.50 (4.80) [30]
	Too High	16.93 (4.44) [29]	17.21 (3.51) [14]	14.52 (3.17) [21]
Current Scale	Appropriate	27.20 (8.71) [71]	26.33 (9.94) [22]	29.20 (12.49) [28]
	Too High	30.24 (9.05) [29]	29.19 (7.93) [16]	28.77 (8.37) [22]
Cognitive Indolence	Appropriate	18.34 (4.66) [73]	20.96 (17.68) [23]	17.06 (5.82) [28]
	Too High	18.83 (5.09) [29]	18.56 (4.86) [16]	17.68 (4.81) [22]
Assertive	Appropriate	54.35 (13.79) [73]	53.13 (13.59) [23]	49.27 (16.54) [32]
	Too High	54.42 (16.88) [29]	41.00 (23.25) [17]	47.30 (17.95) [22]
Aggressive	Appropriate	9.59 (8.91) [73]	14.28 (13.24) [22]	14.82 (12.86) [31]
	Too High	11.48 (8.66) [29]	17.77 (10.65) [17]	15.06 (9.21) [22]
Passive	Appropriate	31.093 (9.42) [73]	32.67 (8.15) [22]	30.34 (8.58) [31]
	Too High	30.93 (8.85) [29]	29.14 (9.88) [17]	27.77 (9.24) [22]

[a] Mean [b] Standard Deviation [c] N

Copas rate: Two one way ANOVAs were undertaken on Copas rate of the completers, non-completers and non-starters: one for the appropriate category of offenders and the other for the too high group. Kolmogorov-Smirnov tests indicated no concerns in relation to normality and the Levene's test indicated no concerns in relation to homogeneity of variance for either the appropriate (F = 0.00, *p* = 1.00) or the too high group (F = 0.55, *p* = 0.946). No significant differences were found between the completion groups in the appropriate group, F = 0.572, *df* = 2, 163, *p* = 0.565 or in the too high group, F = 1.528, *df* = 2, 103, *p* = 0.222. The means are displayed in table 6.15 below.

Table 6.15: Copas rate by completion and appropriateness groups

	Completers	Non-completers	Non-starters
Appropriate	48.96[a] (11.29[b]) [77[c]]	46.58 (10.20) [24]	47.54 (10.21) [65]
Too High	70.54 (14.45) [32]	73.95 (14.04) [20]	76.06 (14.05) [54]

[a] Mean [b] Standard Deviation [c] N

Total OASys score: Two one way ANOVAs were undertaken on Total OASys score of the completers, non-completers and non-starters: one for the appropriate category of offenders and the other for the too high group. Kolmogorov-Smirnov tests indicated no concerns in relation to normality and the Levene's test indicated no concerns in relation to homogeneity of variance for either the appropriate (F = 0.407, *p* = 0.666) or the too high group (F = 0.436, *p* = 0.648). No significant differences were found between the completion groups in the appropriate group, F = 1.923, *df* = 2, 163, *p* = 0.149. The ANOVA relating to the too high group did, however, indicate significance, F = 5.222, *df* = 2, 103, *p* = 0.007. Post-hoc Scheffe tests indicated significant differences between the non-starters and completers: non-starters had significantly higher Total OASys scores than the completers. The full means are displayed in table 6.16 below.

Table 6.16: Total OASys score by completion and appropriateness groups

	Completers	Non-completers	Non-starters
Appropriate	65.40[a] (26.43[b]) [77[c]]	55.29 (24.60) [24]	66.60 (23.04) [65]
Too High	80.44 (24.54) [32]	82.50 (24.80) [20]	95.72 (21.73) [54]

[a] Mean [b] Standard Deviation [c] N

Figure 6.1 demonstrates this relationship further. The scatterplots within figure 6.1 represent those offenders within the 'too high' group only and plot their OGRS2 scores against their Total OASys scores. Within the non-starter group there are more individuals with Total OASys scores above 100 than within the completer and non-completer plots. Indeed where the Total OASys score is above 125, all offenders within the too high appropriateness category failed to commence the programme to which they were sentenced.

Figure 6.1: Scatterplots of OGRS2 risk score against Total OASys score by completion group (too high group only)

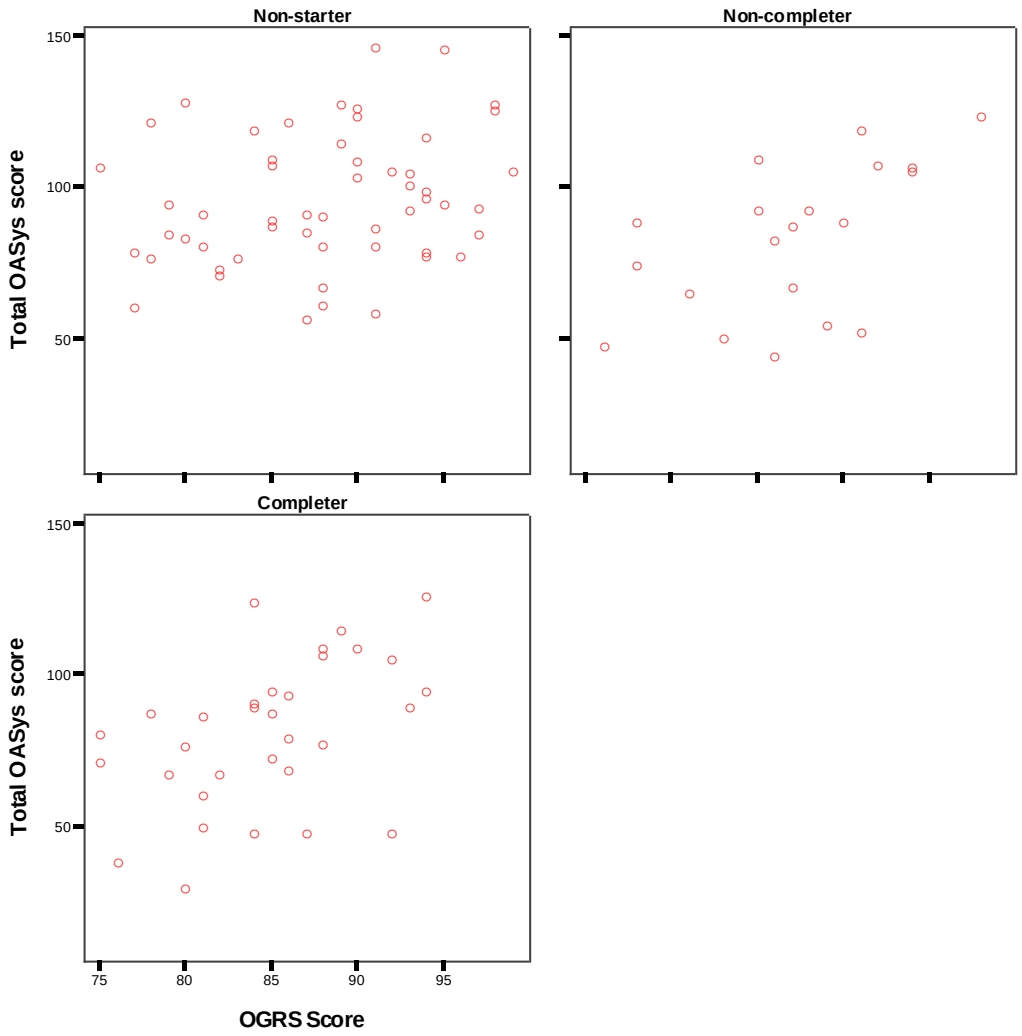
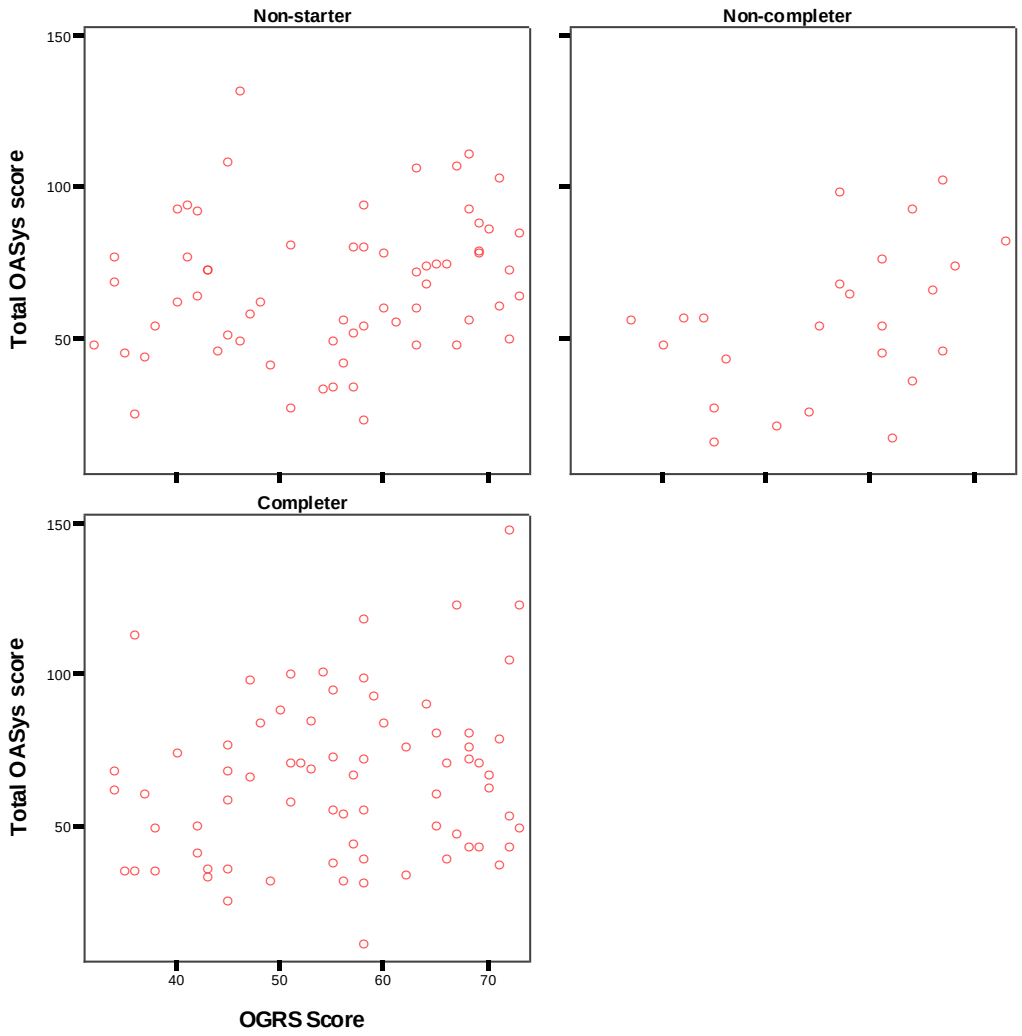


Figure 6.2 presents scatterplots representing the appropriately targeted offenders only. As can be seen and as expected, within this grouping there are fewer offenders with high Total OASys scores. However, where offenders did present with appropriate OGRS2 scores coupled with high Total OASys scores, a higher rate of programme completion was observed amongst this group.

Figure 6.2: Scatterplots of OGRS2 risk score against Total OASys score by completion group (appropriate group only)



Total OASys criminogenic score: Again, two one way ANOVAs were undertaken on Total OASys criminogenic scores of the completers, non-completers and non-starters: one for the appropriate category of offenders and the other for the too high group. Kolmogorov-Smirnov tests indicated no concerns in relation to normality and the Levene's test indicated no concerns in relation to homogeneity of variance for either the appropriate (F = 0.566, *p* = 0.569) or the too high group (F = 1.207, *p* = 0.303). No significant differences were found between the completion groups in the appropriate group, F = 0.458, *df* = 2, 163, *p* = 0.633. The ANOVA relating to the too high group did,

however, indicate significance, F = 5.362, *df* = 2, 103, *p* = 0.006. Post-hoc Scheffe tests indicated significant differences between the non-starters and completers: non-starters had significantly higher Total OASys criminogenic scores than the completers. The full means are displayed in table 6.17 below.

Table 6.17: Total OASys criminogenic score by completion and appropriateness groups

	Completers	Non-completers	Non-starters
Appropriate	28.79[a] (21.48[b]) [77[c]]	24.71 (16.76) [24]	29.08 (19.29) [65]
Too High	33.09 (17.84) [32]	34.25 (15.98) [20]	45.80 (21.14) [54]

[a] Mean [b] Standard Deviation [c] N

Figure 6.3: Scatterplots of OGRS2 risk score against OASys criminogenic need score by completion group (too high group only)

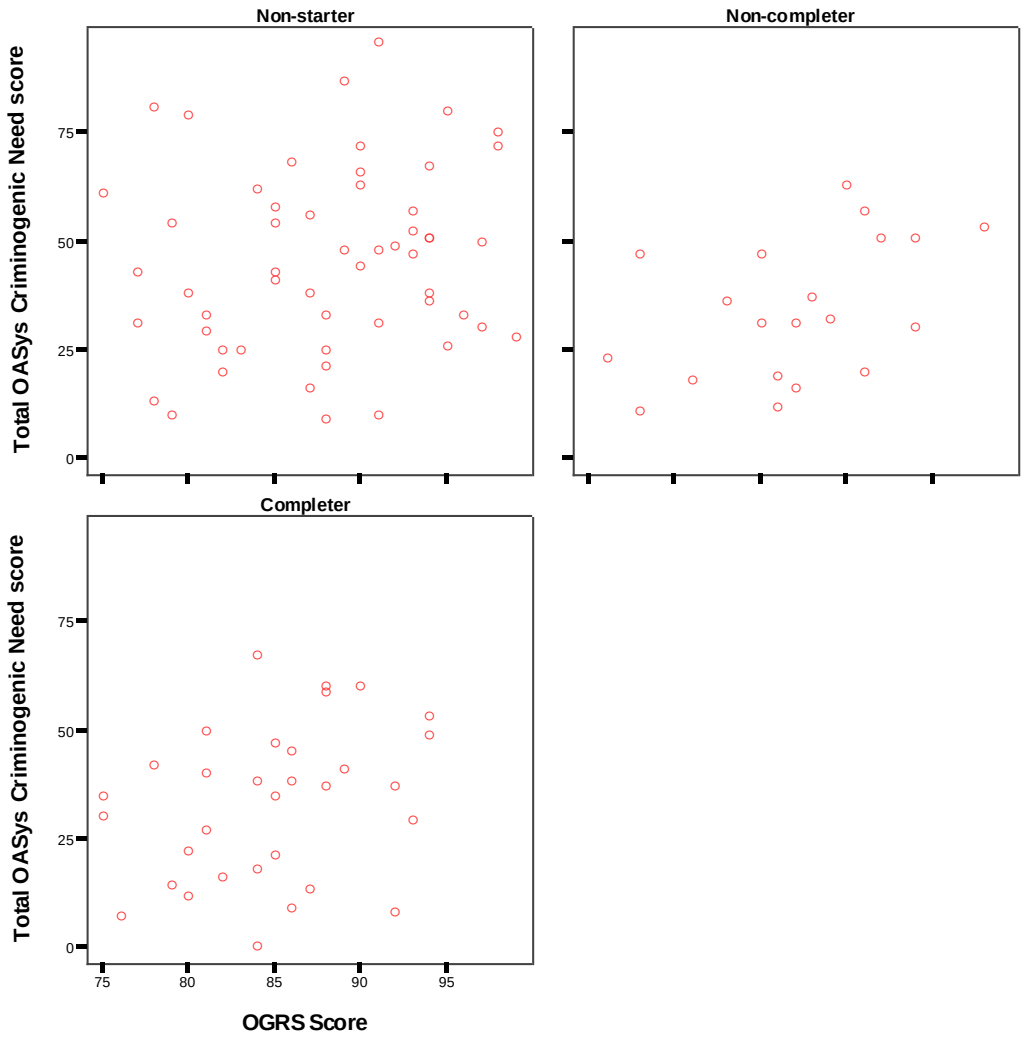
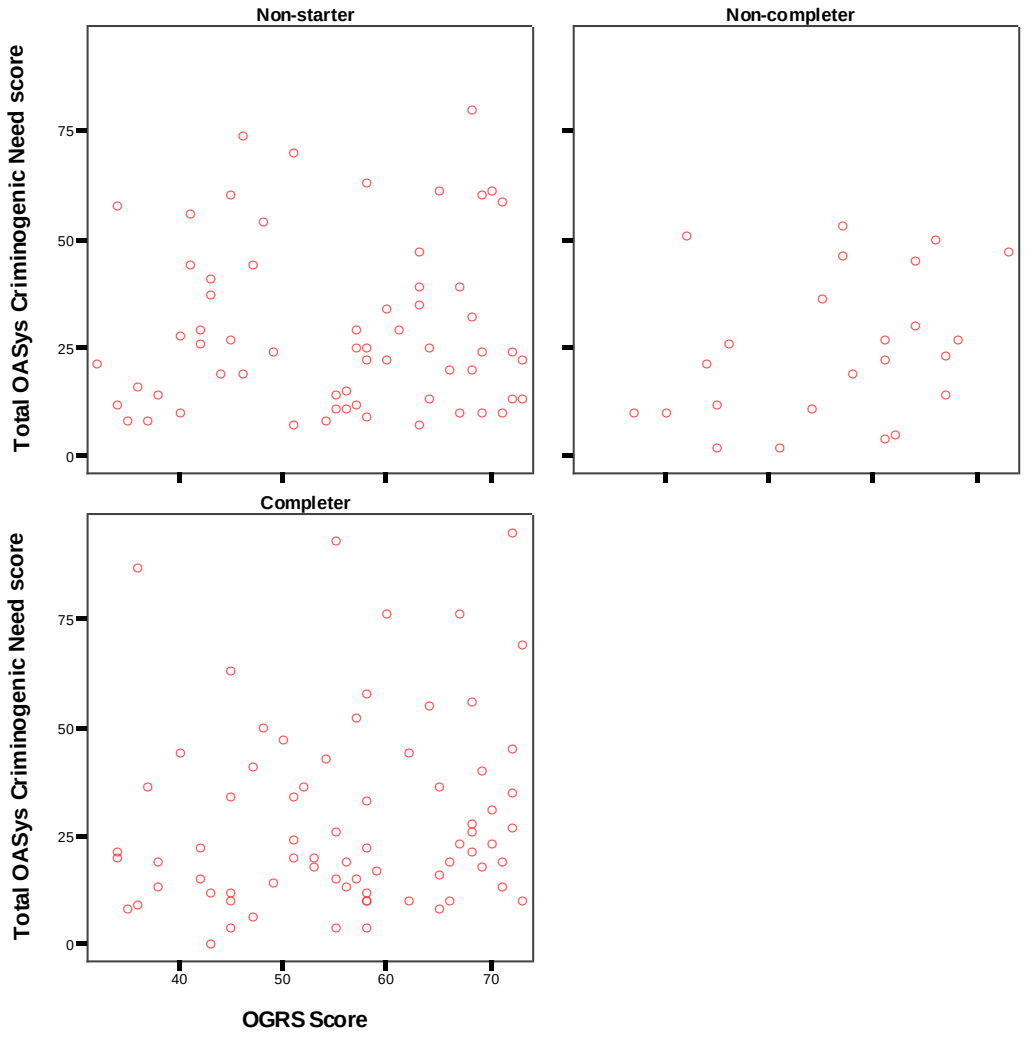


Figure 6.3 demonstrates this relationship in graphical form. The scatterplots represent those offenders within the 'too high' group only and plot their OGRS2 scores against their OASys criminogenic needs scores. Within the completer and non-completer plots there are no individuals with OASys criminogenic need scores above 70. As such, this demonstrates that within the too high category those with OASys criminogenic need scores above 70 all failed to commence the programme to which they were sentenced. Figure 6.4, however, shows that those appropriately targeted

offenders with OASys criminogenic need scores within the high range can and do complete the programme to which they are referred. Finally, another interesting observation which can be taken from these scatterplots is that when an offender within the appropriate targeting group who also has OASys criminogenic need scores above the threshold of 55 commences the programme, they are likely to complete.

Figure 6.4: Scatterplots of OGRS2 risk score against OASys criminogenic need score by completion group (appropriate group only)



Total OASys non-criminogenic score: As above, two one way ANOVAs were undertaken on Copas rate of the completers, non-completers and non-starters: one for

the appropriate category of offenders and the other for the too high group. Kolmogorov-Smirnov tests indicated no concerns in relation to normality and the Levene's test indicated no concerns in relation to homogeneity of variance for either the appropriate (F = 0.342, *p* = 0.711) or the too high group (F = 1.478, *p* = 0.233). No significant differences were found between the completion groups in the appropriate group, F = 0.209, *df* = 2, 163, *p* = 0.811 or in the too high group, F = 0.989, *df* = 2, 103, *p* = 0.375. The means are displayed in table 6.18 below.

Table 6.18: Total OASys non-criminogenic score re by completion and appropriateness groups

	Completers	Non-completers	Non-starters
Appropriate	13.90[a] (10.07[b]) [77[c]]	14.96 (9.38) [24]	14.85 (9.54) [65]
Too High	15.91 (13.11) [32]	16.65 (73.95) [20]	13.17 (9.90) [54]

[a] Mean [b] Standard Deviation [c] N

Discussion

This chapter aimed to determine whether there were any organisational factors associated with completion of, or attrition from, the ETS programme within community settings. This research was able to assess a number of variables, the majority of which are within the influence of the probation service and/or court system.

Within the introduction to this chapter, the potential effect of coercion on engagement with court mandated treatment was discussed. It was assumed that those individuals subject to a suspended sentence order would feel more pressure or coercion from the legal system to complete the programme to which they were allocated. Despite not reaching statistical significance there was some support for this view; half of those on a suspended sentence order compared with just over a third of those on a community order completed their programme. Interestingly, however, the proportion of non-completers did not vary between the two order types: 15% of offenders on both orders commenced but failed to complete the programme. The variability in completion therefore arises from the differences in the number of non-starters between the groups.

Of those on a community order, close to half failed to commence the programme compared with just a third of those on a suspended sentence order. It would seem, therefore, that the increased pressure of a suspended sentence results in more offenders attempting the programme and consequently more completing it.

Of course, it is possible that other factors are associated with the likelihood of receiving a suspended sentence as opposed to community order. Indeed, it is logical that the seriousness of the index offence, and/or the history of offending is likely to be more pronounced amongst those on a suspended sentence. It is possible therefore that the indications observed within these data reflect some other underlying factor rather than the influence of order type. However, given that the variables that are likely to differ between the recipients of the two order types are also those that have so far been found to be associated with higher levels of attrition (rather than lower as observed here), further research is warranted to determine whether there is an association between order type, legally mandated coercion, and attrition from programmes. Such research should investigate the influence of other factors such as risk and criminal history on these relationships and indeed whether the difference between order types does alter the offenders' perceived coercion.

As discussed within the introduction to this chapter, Offender Management National Standards (Ministry of Justice, 2007) dictate that an offender with a requirement to attend an offending behaviour programme should commence the core programme within six weeks (42 days) of their sentence or release. However, for those within the sample for whom this information were available, the mean time from sentence to the commencement of their first programme was over twice this time period (88 days), the maximum being just over a year (399 days). Initially, this finding could be taken as evidence of implementation failure: the programme is not being delivered as it should be. However, the audit criteria for programmes (National Offender Management Service, 2007) state that delays in commencement are permitted if the offender requires motivational or other structured work prior to starting on the programme. Indeed, Turner's (2006) research within the West Yorkshire Probation Area found offenders waiting even longer than this to comment their programmes: completers waited, on average, 121 days whilst the non-completers waited 149 days.

In contrast to the results of Turner (2006), the analyses of the completer, non-completer and non-starters groups revealed there to be no significant differences between the groups' length of time between order and programme commencement. Of course, this analysis did not include those to whom a programme was never allocated: these individuals inevitably became programme non-starters and hence no such data is available for them. What this analysis did reveal, however, was that two-thirds of non-starters were never allocated to a programme. The reasons for this are not known and are likely to reflect a multitude of scenarios. However, it is conceivable that these individuals became classified as non-starters not as a result of their own choice but because of reasons, personal or organisational, which meant that a programme was never *available* to them. For example, an offender could be employed and hence not able to attend, they could have been unsuitably referred to an offending behaviour programme, or the organisation may not have had a programme available for them. Further prospective research should look in more detail at the influence of time from sentencing to the commencement of an available programme with particular attention paid to those who were never allocated to a programme and why this situation may have arisen.

The analysis of the appropriateness of targeting based on OGRS2 scores returned a similar pattern of results to that presented within chapter three. Relative to the non-completer and non-starter groups, there was found to be a larger proportion of offenders with OGRS2 scores below the lower limit of the criterion within the completer group. Additionally, the data showed that the non-starters and non-completer groups have a greater proportion of offenders with OGRS2 score above the upper limit of the targeting criteria than the completer group. Given the findings of Palmer et al. (2008) and Palmer at al. (2009) that attest to no significant differences in reconviction outcomes between the programme completers in the 'too low' category and appropriate comparisons, coupled with the seemingly detrimental outcomes of 'too low' non-completers, it is of concern that probation areas and the courts still persist in referring offenders with OGRS2 score below the targeting criteria to general offending behaviour programmes.

The completion rate amongst those in the 'too high' category was slightly higher than that reported within chapter three (30% as opposed to 26%) which might indicate

better selection of these offenders for programmes or an improved strategy on behalf of the probation service to engage these high risk individuals within treatment. However, the percentage of non-starters within this sample would perhaps argue against such hypotheses: the proportion of offenders in the 'too high' group who fail to attend any programme sessions has increased from 44.5% to just over 50%. As such, the overall dropout rate of 'too high' offenders has remained relatively constant at approximately 70%. Given that the outcomes of those programme completers within the 'too high' category demonstrate large gains (Palmer et al., 2008; Palmer et al., 2009), it is for probation areas to determine whether the 70% attrition rate is outweighed by the treatment gains of the 30% of programme completers.

In an effort to understand more about the completers, non-completers and non-starters within the different appropriateness categories, a series of analyses were undertaken to assess for differences between these groups within the 'appropriate' and 'too high'[23] OGRS2 targeting categories. Within the 'too high' group, offenders with high numbers of criminogenic needs were most likely to fail to commence the general offending behaviour programme. According to Andrews and Bonta (2006) these are the very individuals who are most in need of intervention: they are at a high risk of reconviction and they have high numbers of identified criminogenic needs. Despite this, the chances of gaining and maintaining their engagement in the programme are low. Indeed within the 'too high' category no offenders with criminogenic needs scores above 70 commenced the programme. Within the 'appropriate' category, however, offenders with such scores did complete the programme.

Within the 'appropriate' group, the factors most associated with programme non-completion were the age of the offender – non-completers were younger than completers and non-starters - and their lack of recognition of the impact of their offending on the victims of their offences. As such, it would seem that those offenders who commence a programme who are either young and/or deny the impact of their offence require more support to complete their programme. Additionally and also

[23] Unfortunately, due to low numbers, such analyses were not possible within the 'too low' category. However, the findings of Palmer et al (2008; 2009) would suggest that these offenders should not be sentenced to a general offending behaviour programme as completers do not show gains in terms of their reconviction outcomes and non-completers perform worse than expected. As such, this type of analysis with the 'too low' group would, in any case, be futile.

within the 'appropriate' group, programme dropouts had lower Racial Intimacy scores than programme completers. As discussed within chapter four, low scores on this variable represent more negative attitudes towards racial diversity. Probation areas should therefore consider ways in which they can target resources towards these individuals so as to increase their likelihood of programme completion.

Another factor which might conceivably impact on programme start and/or completion is the time of day at which the programme is scheduled. Within the evaluated probation area and in order to make programmes available to the widest proportion of offenders, the ETS programme is delivered on both a daytime and an evening schedule. The daytime sessions are delivered from mid-morning until lunchtime and the evening sessions commence at 6.30pm. Research has shown that young adults (aged 18 to 22) report that their optimal time of day is either in the evening or are neutral in this regard (May, Hasher, & Stoltzfus, 1993); only 6% of the sample reported a preference for the morning. Further, research has shown that these subjective assessments of optimal times of day are related to measures of cognitive functioning (Bodenhausen, 1990; Rahhal, Abendroth, & Hasher, 1996, cited in Ryan, Hatfield, & Hofstetter, 2002), and memory (May et al., 1993; Petros, Beckwith, & Anderson, 1990). It is possible therefore that the time of day of the programme could impact on engagement with the material and hence could be linked with attrition and programme outcomes.

In viewing the proportions of completers, non-completers, and non-starters within each time slot, however, it can be seen that there is very little variation in attrition across the two delivery times. There was a slight indication that those on the day programme were more likely to complete and less likely to non-complete than those on the evening programme, however, there was no significant association between group and time of programme delivery.

The speed at which a programme is delivered is one of the considerations that comprise the programme accreditation criteria relating to dosage and sequencing. In line with the responsivity principle (Andrews & Bonta, 1994), programmes should be delivered to offenders' in a manner which is responsive to their learning needs. As such, it is possible that the pace at which a programme is delivered is important to the success of that programme: there must be adequate time between sessions for programme

attendees to consolidate their learning but too long a gap could mean that previous session contents are forgotten and hence the impact of that session, and how future sessions might build upon it, is lost.

While it is not possible to assess the impact of the different pace of delivery on programme outcomes from the data available within this study, it has been possible to assess whether there is a relationship between pace and programme completion. The vast majority of offenders (over four in five) were allocated to a programme which ran twice a week over ten weeks. There was little impact of pace on programme commencement but, once the programme had commenced, programme completion was more likely on the thrice weekly programme. This finding, however, should be interpreted cautiously for two reasons: first, there was no statistical significance in relation to this distribution and hence can only be interpreted, at most, as a trend in the data ($\Phi_c = 0.10$); and second, it is not possible from these data to tease out whether the trend is a function of programme pace or of those allocated to the different paces. It could be that the reasons for allocation to different rates of delivery are responsible for the trend rather than the rate itself. Further research should assess this in more detail.

The final organisational factor assessed in relation to programme completion was the number of attempts that each individual had at attending and completing a programme. At present, the Probation Area evaluated allows a maximum of three attempts at programme completion (in rare cases, a further attempt might be allowed by the programme manager; there were eight such examples of this within the sample). This research found that non-starters have the fewest attempts at programme completion: this is perhaps an obvious finding for three reasons: first, some non-starters were never referred to a programme in the first place; second, some reasons for failing to commence a programme might also prevent them from having another attempt, for example, if they are held in custody or if they have been unsuitably referred; and third, if they are generally non-compliant with their order they may be subject to breach proceedings and hence do not wish to comply until these are concluded.

In relation to completers and non-completers, completers had significantly fewer attempts at programme completion. Notwithstanding the potential effects that increased numbers of failed attempts may have on treatment outcomes (to the author's knowledge there is currently no research investigating this issue), the data within table

6.3 shows that the permitting of further attempts is worthwhile. At attempt one, only 29% of the sample completed the programme. This rose to 43% after the further attempts were allowed. Indeed, the completion rate increased at each attempt from the initial 29% at attempt one to 50% at attempt four. From these data, therefore, it is recommended that the policy to allow further attempts at programme completion remains. Further research is needed, however, to determine whether the number of attempts a person has at a programme has an effect on the impact of the programme on the individual. This could be tested against intermediate factors, such as behavioural measures and psychometric scores, in addition to reconviction or recidivism outcomes.

In conclusion, this chapter has investigated process or organisational factors to determine whether attrition is influenced by factors external to the individual. Although trends within the data have been highlighted and discussed, very few of the factors considered within this chapter actually returned statistically significant results. Not surprisingly given the research presented in chapter three, there was a high level of inappropriate referrals to the ETS programme which resulted in high completions rates amongst the 'too low' category and high non-starter rates amongst the 'too high' category of offenders. In addition, it was observed that permitting further attempts on the programme did increase the completion rate substantially. Finally, probation officers ratings of the offenders' levels of engagement on the programme were found to predict attrition from the ETS programme. It would seem, therefore, that some process or organisational factors do impact on the likelihood of programme completion and hence the responsibility for this cannot be placed solely within the offender. Such findings reported here lend support to the 'readiness to change' theory of Ward and colleagues (Ward et al., 2004); not only does the individual need to be ready to address their offending behaviour, the situational aspects of treatment also need to be conducive to, and supportive of, the change process.

Chapter Seven

An analysis of the reasons for attrition

Introduction

The previous chapters within this thesis have assessed differences in the characteristics of non-starters, non-completers and completers of offending behaviour programmes within the community correctional service of England and Wales. In addition, the previous chapter has investigated whether these groups are processed differently by the organisation and hence whether process factors are associated with attrition. The preceding research has therefore been concerned with the correlates of attrition: those factors that are associated with or co-occur with programme dropout or completion. However, in order to reduce attrition rates it is also necessary understand the reasons why an offender fails to attend or complete the intervention to which he or she has been sentenced.

Stephens (2003) has argued that "For reasons for non-completion to inform change, they need to point to causal factors that can be addressed" (p. 234). In reality, however, the information held within probation systems does not always indicate causal factors. As Stephens and Turner (2004) state: "many 'reasons' for drop out will not be recorded in an easy to interpret way in case records, and some may not be recorded at all" (p. 18). In addition, Stephens argues that programme attrition may occur due to a series of events rather than one single reason and that this information cannot be represented easily within case files. Despite these issues, Stephens and Turner were able to code case documentation, supplemented with verbal information from case managers and administrative staff, into the action-orientated typology of attrition as proposed by Stephens in her earlier paper (2003).

As outlined within the literature review, Stephens proposes that a hierarchical method of questioning can determine the most relevant reason for failure to start a programme (Stephens' method of classification is concerned with pre-programme attrition only). The author argues that such a methodology can help clear some of the haze around why an offender did not attend the planned intervention. The typology by Stephens contains three types of attrition: first, those who could not attend, second, those who could but should not attend, and third, those who could, should, but did not attend their intervention.

The first of these, those who could not attend their programme, outlines those reasons that make it impossible for an offender to attend a programme. Stephens argues that if such a reason, or a 'type one factor', is found then the questioning can cease at this point: "what Type 1 factors should all have in common is an implication that no further action either could or should be taken" (p. 236). The sorts of factors that would result in the offender not being able to attend the programme are: being in custody, having transferred out of the area, death, hospitalisation, or whereabouts unknown.

Type two factors, or those relating to offenders who could but should not attend the programme, focus on the suitability of the offender for the intervention. As outlined earlier, accredited offending behaviour programmes are required to outline clear targeting criteria relevant to attendance on the programme (Lipton et al., 2000). As such, each programme specifies the criteria by which offenders should be selected. In addition, the National Management Manual (2000) indicates that offenders should be excluded from attending an offending behaviour programme if they do not demonstrate a need for it, they have mental health problems, they have other factors which could affect attendance, for example, drug dependency, or they have low IQ levels. Stephens (2003) also extends type two factors to include complete denial of the offence, and being unable to cope with learning within a group setting. As Stephens outlines "What the factors should have in common is that they point to the improvement of assessment processes to ensure that offenders are allocated correctly from the start" (p. 236).

Type three factors are those which represent that attrition whereby the offender could, should, but did not attend the intervention to which they were sentenced. Stephens claims that this type of attrition "represents the real leaks in the system through which offenders who are available, eligible and suitable (who *could* and *should* attend) are missing opportunities for access to rehabilitative experiences" (p.236). Stephens initially proposed that type three attrition be broken into three broad categories: those relating to programme design, to local organisation, or to offender motivation. However, the proposed sub-categories were derived through general observation by the researcher and hence were not evidence-based. As such when classification was attempted by Stephens and Turner the categories did not match completely with the data and hence they were modified. The resultant categories were:

3A – those for whom a group place was effectively not available, 3B – those for whom a group place was available but the offender was in breach of their order and 3C – those for whom a group place was available, the offender was not in breach of their order but they did not attend. As such, these are hierarchically situated and provide indications as to at which point in the process intervention should be targeted to increase attendance.

Stephens and Turner (2004) have used this conceptual model within West Yorkshire Probation Area to classify offenders who failed to commence an accredited programme within three months of their sentence date. They reported that 30% of the sample had started the programme, eight percent were not able to attend, six percent were able but shouldn't attend, and the remaining 56% could, should but didn't attend the programme to which they were sentenced. The statistics relating to the general offending behaviour programme, ETS, did not differ hugely from the collective: 25% had commenced the programme, seven percent could not, five percent should not, and 63% could, should but didn't attend the ETS programme. As such, the vast majority of attrition was classified at type three: three in four of all those who dropped out of their programme were available, suitable and eligible to attend. Of these 44%, did not attend the ETS programme due to places on the programme not being available to them. The main reason for this, as discovered by Stephens and Turner, was the scheduling of programmes that were not flexible enough to account for offender employment patterns.

An additional 44% of type three offenders were classified as 3B: they were in breach of their order prior to commencement of the programme and hence had not started the programme as a result of this. Further investigation of these cases, however, showed that this group was not homogeneous: some had never been seen by their case manager whilst others had been subject to breach proceedings but were now compliant with their order. As such, it might be expected that this latter group might commence a later accredited programme should they be referred.

Finally, the remaining 12% of type three offenders were eligible and suitable to attend, a group place was available, they were not in breach of their order but they had still failed to commence the programme. Stephens and Turner state that in most of these cases there was no indication within the case record as to why this attrition had

occurred. Further research is therefore required with this group of offenders to determine why they did not commence their intervention.

What this research shows is that, despite the view implicit in research which assesses only the characteristics of dropouts that attrition is the consequence of factors internal to the offender, a large proportion of attrition results directly due to factors outside of the volitional control of the offender. For example, 30% of offenders sentenced to attend an accredited programme were either unsuitable for programme work or did not have a place on a programme available to them that matched their particular circumstances within three months. Whilst it is not necessarily the case that all these offenders would, if given the chance, have completed the programme, such information does question the organisational impact on attrition and also the validity of the combination of such offenders with those who chose not to attend within attrition research.

The research within this chapter will therefore attempt to utilise the classification scheme of Stephens (2003) and Stephens and Turner (2004) to determine the reasons for attrition from the ETS programme within one probation area in England and Wales. However, the classification system proposed in these papers will be modified slightly. First, in taking offenders who had not commenced within three months of their order, Stephens and Turner have presented an incomplete picture of events within the probation area. Given the findings in chapter six above relating to the length of time between order and programme commencement (the average was just short of 90 days), it is possible that the three month cut off utilised within this research would have prematurely classified the offenders: it could well be that some offenders classified in this research as non-starters would ultimately attend and even complete the programme to which they were sentenced. As such, the research within this chapter will utilise retrospective data relating to offenders whose orders have been completed (either successfully or unsuccessfully). As such, the final status of these offenders in relation to their offending behaviour programme is known. As, however, it is known (and as is presented in chapter six) that many offenders do not manage to attend or complete ETS at their first attempt the analysis will be undertaken using first, the offenders first attempt at the programme, and second, their final status. Breaking the

data down in this manner will allow investigation and comparison across the two attempts.

Second, the classification system devised by Stephens and utilised by Stephens and Turner investigated the reasons for offenders failing to *commence* their intervention. However, there is no reason why this system cannot be extended to look at in-programme attrition. As such, this research will classify both the non-starters and the non-completers to determine the reasons for dropout from the ETS programme.

Third, the Stephens and Turner's research utilised a sample of just 166 offenders across four accredited programmes: the ETS data related only to 55 participants of whom a quarter commenced the programme. It is possible, indeed likely, that the offence specific programmes, such as substance use or domestic violence programmes, included in the research of Stephens and Turner have different factors contributing to attrition. Indeed, the authors reported differences between the programmes in relation, for example, to the number of offenders who were deemed unsuitable due to denial of the offence. In order to improve on the Stephens and Turner paper, therefore, this chapter will utilise a sample of nearly 300 offenders sentenced to the ETS programme within one probation area to determine the reasons for dropout.

Method

Design:

This chapter aimed to classify the non-starters and non-completers of the ETS programme into the action-oriented typology proposed by Stephens (2003). As such, the non-starters and non-completers were classified into the type one (those who could not attend), type two (those who could attend but shouldn't attend) and type three (those who could and should but didn't attend) categories as outlined in Stephens (2003) and Stephens and Turner (2004). Furthermore and where appropriate, a three (attrition groups: type one, type two, type three) by two (dropout group: non-starter, non-completer) between-groups design was utilised to test for significant differences between the groups on those variables that have been seen in previous chapters to differ significantly between the completers, non-completers, and non-starters. This will allow more in-depth examination of the dropout groups to determine whether further subsets of offenders are present within the non-starter and non-completer groups.

Participants:

The participants utilised within chapters four, five, and six were once again utilised within this chapter. However, these analyses referred only to the programme dropouts and hence the completers were removed from the analysis. Initially, the data relating to the offenders' first attempt on the ETS programme was utilised. Of the 293 participants within the overall sample, 85 (29.01%) completed the ETS programme on their first attempt. These individuals were therefore removed from the initial analysis leaving 208 programme dropouts. The second set of analyses referred to the final completion/dropout status and as such the 126 offenders who completed the ETS programme were removed from the analysis, leaving 167 programme dropouts.

Measures and procedure:

As with the previous three chapters, the data was extracted from the three probation databases (CRAMS, IAPS, and OASys) within the local probation area. The classification of offenders into the non-starter and non-completer groups was undertaken as previously described. The classification of offenders into the attrition groups (type one, type two, and type three) was undertaken utilising information from CRAMS and IAPS. For each offender, the contact notes (held within CRAMS) were read and a reason for non-attendance at the ETS progamme was determined. Where possible, these data were cross-checked against the IAPS entry relating to the reason for the offenders' suspension from the programme to which they allocated. This was not possible, however, in cases where the offender was not allocated to a programme (all of these were non-starters, but not all non-starters were not allocated): an entry is only created in IAPS once an offender is referred to a particular run of the programme.

Once this information had been collected from CRAMS and cross-checked against the IAPS entry for each individual, classification into Stephens' typology was undertaken. Using Stephens' (2003) list of type one, type two, and type three reasons for attrition, the offenders were coded appropriately. As outlined above, rather than only code the non-starters as done by Stephens and Turner (2004), the non-completers of programmes were also coded into the typology.

Results

Status at first attempt:

The analysis was initially conducted using each offender's first attempt at the ETS programme. Of the 293 participants within the overall sample, 85 (29.01%) were completers of the ETS programme. These individuals were therefore removed from the analysis. Of the remaining 208 participants, 19 (9.13%) were, on the basis of their first programme attempt, classified as Type One (offender not available), 19 (9.13%) were classified as Type Two (offender not eligible and/or not suitable), and the remaining 170 (81.73%) were classified as Type Three ('true' attrition).

Type one reasons: Table 7.1 presents the reasons provided for the non-attendance of the type one individuals.

Table 7.1: Reasons for type one non-attendance

	N (% of Type one)	N of Non-starters (%)	N of Non-completers (%)
Offender in custody	7 (36.84)	5 (33.33)	2 (50.00)
Offender at court/breach appearance	7 (36.84)	5 (33.33)	2 (50.00)
Offender moved out of area	5 (26.32)	5 (33.33)	-
Total	19 (100.00)	15 (100.00)	4 (100.00)

As can be seen, the majority of type one individuals were not able to attend the programme due to ongoing action relating to the administration of justice: 66.66% of the sample were either in custody or were unable to attend due to a court appearance which took precedence over programme attendance. Of the non-starters, a third moved out of the area and hence may have commenced the programme with their new probation area. Of the non-completers, half failed to complete due to being in custody whilst the remainder were in breach or in court.

Type two reasons: Table 7.2 presents the reasons for programme non-attendance of the type two individuals.

Table 7.2: Reasons for type two non-attendance

	N (% of Type two)	N of Non-starters (%)	N of Non-completers (%)
Programme unsuitable – group work	6 (31.58)	4 (25.00)	2 (66.66)
Requirement deleted	3 (15.779)	3 (18.75)	-
Programme unsuitable – mental health	2 (10.53)	2 (12.50)	-
Programme unsuitable – chaotic	2 (10.53)	2 (12.50)	-
Programme unsuitable – literacy	2 (10.53)	2 (12.50)	-
Programme unsuitable – comprehension	1 (5.26)	-	1 (33.33)
Offender previously completed ETS	1 (5.26)	1 (6.25)	-
Offender allocated to an alternative prog.	1 (5.26)	1 (6.25)	-
Prog. never a requirement on the order	1 (5.26)	1 (6.25)	-
Total	19 (100.00)	16 (100.00)	3 (100.00)

All of the type two non-completers failed to complete due to the programme being unsuitable for them: two were not suitable for group work and the remaining person did not comprehend the programme. Of the non-starters, 62.50% were not suitable for the programme (although this may be higher as it may be possible that the three offenders who had the programme requirement deleted from their order were also unsuitable for the programme). In addition, 18.75% of non-starters were referred to the programme due to administrative errors: one had previously completed ETS, another was allocated to an alternative programme (although it is not possible to determine from the files why this occurred), and another never actually had ETS as a requirement on their order.

Type three reasons: Table 7.3 presents the reasons provided for the non-attendance of the type three ('true' attrition) individuals. As can be seen below, this attrition category accounted for the majority of programme dropout. Further classification of this category into Stephens and Turner's (2004) 3A, 3B, and 3C groups

demonstrated that a significant proportion (17.65%) of such attrition was influenced by organisational factors: no programme place was available or those available did not meet the needs of the sample. A further third of type three attrition was due to the offenders either being in breach of appointments other than those relating to ETS or committing a further offence. The largest proportion of type three attrition, however, was type 3C and occurred when the offender was apparently compliant with other aspects of their order but did not attend their ETS sessions. The reasons given, for example, were that the offender was ill, they were breached for behavioural reasons, they were confused about the dates of the programme sessions, or, as in the majority of cases, they just failed to attend and no reason was provided.

Table 7.3: Reasons for type three non-attendance

	N (% of Type three)	N of Non-starters (%)	N of Non-completers (%)
3A reasons: a group place was effectively not available:			
Employment	18 (10.84)	12 (10.71)	6 (10.34)
Child care arrangements	3 (1.81)	2 (1.79)	1 (1.72)
Attended but curfew not amended	2 (1.20)	2 (1.79)	-
Accommodation problems	2 (1.20)	2 (1.79)	-
Prog. suspended due to low numbers	2 (9.09)	-	2 (3.45)
Prog. not available – lone female	1 (5.00)	1 (0.89)	-
Prog. conflicted with other order reqs	1 (4.54)	1 (0.89)	-
Order expired	1 (0.60)	-	1 (1.72)
3A Subtotal	30 (17.65)	20 (17.86)	10 (17.24)
3B reasons: in breach of other appointments:			
Failed to comply: other aspects	33 (19.88)	26 (23.21)	7 (12.07)
Further offence	27 (16.26)	25 (22.32)	2 (1.72)
3B Subtotal	60 (35.29)	51 (45.54)	9 (15.52)

Table 7.3: Reasons for type three non-attendance (cont.)

	N (% of Type three)	N of Non-starters (%)	N of Non-completers (%)
3C reasons: apparently compliant with other aspects:			
Non-attendance	47 (28.31)	28 (25.00)	19 (32.76)
Illness	15 (9.04)	6 (5.36)	9 (15.52)
Late to sessions	7 (4.22)	2 (1.79)	5 (8.62)
Breached behavioural requirement	3 (1.81)	-	3 (5.17)
Confused about dates	2 (1.20)	1 (0.89)	1 (1.72)
Not motivated so not allocated by OM	1 (0.60)	1 (0.89)	-
Would not commit to dosage	1 (0.60)	1 (0.89)	-
Conflicted with signing on	1 (0.60)	1 (0.89)	-
Illness (self induced: alcohol)	1 (0.60)	-	1 (1.72)
Booked holiday	1 (0.60)	-	1 (1.72)
Disliked the tutor	1 (0.60)	1 (0.89)	-
3C Subtotal	80 (47.06)	41 (36.61)	39 (67.24)
Total	170 (100.00)	112 (100.00)	58 (100.00)

Status at final attempt:

The analysis was re-run using each offender's final attempt at the ETS programme. Where the offender had more than one attempt at the ETS programme, the reason for non-attendance relating to their final attempt was utilised within this analysis. As such, this analysis represents the offenders' final classification (completer, non-completer, non-starter) and hence matches with the classifications used within chapters four to six. Of the 293 participants within the overall sample, 126 (43.00%) were completers of the ETS programme. These individuals were therefore removed from the analysis. Of the remaining 167 participants, 21 (12.57%) were classified as Type One (offender not available), 20 (11.98%) were classified as Type Two (offender not eligible and/or not suitable), and the remaining 126 (75.45%) were classified as Type Three ('true' attrition).

Type one reasons: Table 7.4 presents the reasons provided for the non-attendance of the type one individuals.

Table 7.4: Reasons for Type One non-attendance

	N (% of Type one)	N of Non-starters (%)	N of Non-completers (%)
Offender in custody	10 (47.62)	5 (38.46)	5 (62.50)
Offender moved out of area	6 (28.57)	6 (46.15)	-
Offender at court/breach appearance	5 (23.81)	2 (15.38)	3 (37.50)
Total	21 (100.00)	13 (100.00)	8 (100.00)

The table above shows that the type one programme non-completers were not able to complete the programme either due to being taken into custody during the programme or due to breach proceeding and/or being taken back to court. Unfortunately the data was not detailed enough to determine why the non-completers were taken into custody or breached.

A significant proportion of non-starters were also not able to attend the programme due to being in custody or in breach/at court (50.00%). An additional six offenders were not able to commence the programme due to transference out of the locality.

Type two reasons: Table 7.5 presents the reasons provided within the system for programme non-attendance of the type two individuals.

Table 7.5: Reasons for Type Two non-attendance

	N (% of Type two)	N of Non-starters (%)	N of Non-completers (%)
Programme unsuitable – group work	6 (30.00)	4 (23.53)	2 (66.66)
Requirement deleted	4 (20.00)	4 (23.53)	-
Programme unsuitable – mental health	2 (10.00)	2 (11.76)	-
Programme unsuitable – chaotic	2 (10.00)	2 (11.76)	-
Programme unsuitable – literacy	2 (10.00)	2 (11.76)	-
Programme unsuitable – comprehension	1 (5.00)	-	1 (33.33)
Offender previously completed ETS	1 (5.00)	1 (5.88)	-
Offender allocated to an alternative prog.	1 (5.00)	1 (5.88)	-
Prog. never a requirement on the order	1 (5.00)	1 (5.88)	-
Total	20 (100.00)	17 (100.00)	3 (100.00)

As with the first attempt statistics, all of the type two non-completers failed to complete the programme due to unsuitability issues: two were not suitable for groupwork and one lacked the ability to comprehend the programme contents. However, these issues were not recognised until after the offender commenced the programme.

Of the non-starters, almost two thirds (62.50%) were also assessed as unsuitable following receipt of the order to attend the programme. A further three non-starters had their requirement deleted (although it was not possible from the data available to determine why these requirements were deleted). Furthermore three non-starters were administrative non-starters – one had previously completed ETS, one was allocated to a

programme other than ETS after additional assessment and the final offender never had the ETS requirement on their order.

Type three reasons: Table 7.6 presents the reasons provided for the non-attendance of the type three ('true' attrition) individuals.

Table 7.6: Reasons for Type three non-attendance

	N (% of Type three)	N of Non-starters (%)	N of Non-completers (%)
3A reasons: a group place was effectively not available:			
Employment	16 (12.70)	11 (11.96)	5 (14.71)
Child care arrangements	2 (1.59)	2 (2.17)	-
Prog. conflicted with other order reqs	1 (0.79)	1 (1.09)	-
Accommodation problems	1 (0.79)	1 (1.09)	-
Order expired	1 (0.79)	-	1 (2.94)
3A Subtotal	21 (16.67)	15 (16.03)	6 (17.65)
3B reasons: in breach of other appointments:			
Further offence	30 (23.81)	26 (28.28)	4 (11.76)
Failed to comply: other aspects	28 (22.22)	26 (28.26)	2 (5.88)
3B Subtotal	58 (46.03)	52 (56.52)	6 (17.65)
3C reasons: apparently compliant with other aspects:			
Non-attendance	33 (26.19)	22 (23.91)	11 (32.35)
Breached behavioural requirement	6 (4.76)	-	6 (17.65)
Illness	2 (1.59)	-	2 (5.88)
Confused about dates	2 (1.59)	-	2 (5.88)
Late to sessions	1 (0.79)	1 (1.09)	-
Not motivated so not allocated by OM	1 (0.79)	1 (1.09)	-
Would not commit to dosage	1 (0.79)	1 (1.09)	-
Booked holiday	1 (0.79)	-	1 (2.94)
3C Subtotal	47 (37.30)	25 (27.17)	22 (64.71)
Total	126 (100.00)	92 (100.00)	34 (100.00)

As with the data relating to the first attempt at the programme, the type three attrition category accounted for the majority of programme dropout. Comparative to the first attempt data, a similar proportion of type three attrition was due to a lack of appropriate programme places (first attempt, 17.65%; final attempt, 16.67%). However, the proportions of 3B and 3C attrition differed from those seen on the first attempt. Almost half of type three attrition (compared with just a third at first attempt) was due to the offender either being subject to breach proceedings or having committed a further offence. Finally, just over a third of attrition (compared with a half at first attempt) occurred when the offender was apparently compliant with other aspects of their order but failed to attend the ETS programme.

Correlates of type one, two and three attrition

The analyses within chapters four, five and six were examined to determine which variables differed significantly between the completers, non-completers, and non-starters of the ETS programme. These data were then subject to further analyses to determine whether there were any differences between the final status attrition groups in relation to these variables. The variables involved within these analyses therefore were: age, OGRS2 score, pre-programme racial intimacy and aggressiveness scores, previous breach, number of previous convictions (total and prior to age 18), age at first conviction, Copas rate, OASys total score, OASys criminogenic need score, number of attempts at ETS on this order and OGRS2 appropriateness.

Initially, the aim was to conduct a series of three (type one, type two, type three) by two (non-starters, non-completers) ANOVAs on all of the continuous variables. However, some of the variables failed to meet the parametric assumption of normality and hence non-parametric alternatives had to be utilised for these variables. Where this was necessary, Kruskal Wallis was computed initially for the full sample with attrition type (type one, type two, type three) as the independent variable. Following this, the sample was split into non-starters and non-completers and the analyses re-run i.e. one Kruskal-Wallis test was undertaken using the non-completer sample and another with the non-starter sample. Where significant differences were found, Mann Whitney U tests were utilised to assess the pairwise comparisons for significant differences.

The categorical variables were analysed using chi-square as appropriate, first on the overall sample and then separately on the non-starter and non-completer samples.

Age: Table 7.7 presents the mean ages of the three attrition groups in total and by non-starter and non-completer groups. Kolmogorov-Smirnov tests revealed that this variable was non-normally distributed amongst one of the attrition groups (type three, Kolmogorov-Smirnov Z (*K-S Z*) = 2.036, $p = 0.001$) and hence non-parametric tests were conducted to test for differences in age between the attrition groups. The Kruskal-Wallis test revealed a significant difference between the three groups using the overall sample (χ^2 (2, $N = 167$) = 10.464, $p = 0.005$, $\Phi_c = 0.25$). Pairwise comparisons utilising Mann-Whitney U tests revealed significant differences between the type one and type two ($U = 115.50$, $p = 0.013$, r = 0.38) and type two and type three groups ($U = 701.50$, $p = 0.001$, r = 0.26). No differences were found between the type one and type three groups ($U = 1311.00$, $p = 0.947$, r = 0.01). The type two offenders were hence significantly older than the type one and type three offenders.

Table 7.7: Age by Attrition type (type one, type two, type three) and Group (non-starter, non-completer)

	Type one (N = 21)	Type two (N = 20)	Type three (N = 126)
Non-starters	23.15[a] (5.31[b]) [13[c]]	31.00 (8.97) [17]	23.42 (5.62) [92]
Non-completers	24.38 (10.90) [8]	24.00 (6.56) [3]	22.56 (5.97) [34]
Total	23.62 (7.67) [21]	29.95 (8.88) [20]	23.19 (5.70) [126]

[a] Mean [b] Standard Deviation [c] N

Further Kruskal-Wallis tests on the separate groups of non-starters and non-completers revealed that there were significant differences between the attrition groups within the non-starter sample (χ^2 (2, $N = 122$) = 11.052, $p = 0.004$, $\Phi_c = 0.30$) but not amongst the non-completer sample (χ^2 (2, $N = 45$) = 0.266, $p = 0.876$, $\Phi_c = 0.05$).

Pairwise comparisons utilising Mann-Whitney U tests with the non-starter sample revealed significant differences between the type one and type two groups ($U =$

50.50, $p = 0.012$, $r = 0.46$) and the type two and type three groups ($U = 394.50$, $p = 0.001$, $r = 0.31$). No differences were found between the type one and type three groups ($U = 588.50$, $p = 0.926$, r = 0.01). The type two offenders were hence significantly older than both type one and type three offenders.

OGRS2: Table 7.8 presents the mean OGRS2 scores of the three attrition groups in total and by group (non-starter, non-completer). As with the age variable, Kolmogorov-Smirnov tests revealed that this variable was non-normally distributed amongst one of the attrition groups (type three, *K-S Z* = 1.383, $p = 0.044$) and hence non-parametric tests were again used to test for differences between the attrition groups. The Kruskal-Wallis test revealed a significant difference between the three groups using the overall sample (χ^2 (2, $N = 167$) = 11.735, $p = 0.003$, $\Phi_c = 0.27$). Pairwise comparisons utilising Mann-Whitney U tests revealed significant differences between the type one and type two ($U = 100.00$, $p = 0.004$, r = 0.44) and type two and type three groups ($U = 676.50$, $p = 0.001$, r = 0.27). No differences were found between the type one and type three groups ($U = 1290.50$, $p = 0.857$, r = 0.01). The type two offenders were hence significantly lower risk than the type one and type three offenders.

Table 7.8: OGRS2 by Attrition type (type one, type two, type three) and Group (non-starter, non-completer)

	Type one (N = 21)	Type two (N = 20)	Type three (N = 126)
Non-starters	70.46[a] (20.16[b]) [13[c]]	53.88 (19.16) [17]	71.41 (19.61) [92]
Non-completers	74.88 (13.72) [8]	54.67 (10.97) [3]	68.97 (19.87) [34]
Total	72.14 (17.74) [21]	54.00 (17.94) [20]	70.75 (19.63) [126]

[a] Mean [b] Standard Deviation [c] N

Further Kruskal-Wallis tests on the separate groups of non-starters and non-completers revealed, as with the age variable, that there were significant differences between the attrition groups within the non-starter sample (χ^2 (2, $N = 122$) = 9.649, $p = 0.008$, $\Phi_c = 0.28$) but not amongst the non-completer sample (χ^2 (2, $N = 45$) = 2.336, $p =$

0.311, $\Phi_c = 0.23$). Pairwise comparisons utilising Mann-Whitney U tests with the non-starter sample revealed significant differences between the type one and type two (*U* = 61.00, *p* = 0.038, r = 0.38) and type two and type three groups (*U* = 412.00, *p* = 0.002, r = 0.29). No differences were found between the type one and type three groups (*U* = 581.00, *p* = 0.869, r = 0.02). The type two offenders were hence significantly lower risk than the type one and type three offenders.

Pre-programme Racial Intimacy: A three (type one, type two, type three) by two (non-starter, non-completer) between subjects ANOVA was conducted to test for differences between the groups in relation to the pre-programme racial intimacy variable as the Kolmogorov-Smirnov tests revealed no problems with normality and the Levene's Test indicated no problems with homogeneity of variance (F = 0.195, p = 0.963). No significant differences were revealed between the attrition groups (F = 0.108, *df* = 2, 81, *p* = 0.898, r = 0.05) or the dropout groups (F = 1.525, *df* = 1, 81, *p* = 0.220, r = 0.13). The interaction was also non-significant (F = 0.204, *df* = 2, 81, *p* = 0.816, r = 0.07). The means are displayed in table 7.9 below.

Table 7.9: Pre-programme Racial Intimacy by Attrition type (type one, type two, type three) and Group (non-starter, non-completer)

	Type one (N = 9)	Type two (N = 10)	Type three (N = 68)
Non-starters	15.50[a] (4.04[b]) [4[c]]	14.38 (4.53) [8]	15.10 (4.31) [41]
Non-completers	16.00 (3.24) [5]	17.00 (4.24) [2]	17.44 (3.94) [27]
Total	15.78 (3.38) [9]	14.90 (4.38) [10]	16.03 (4.29) [68]

[a] Mean [b] Standard Deviation [c] N

Pre-programme Aggressiveness: A three (type one, type two, type three) by two (non-starter, non-completer) between subjects ANOVA was conducted in relation to the pre-programme aggressiveness variable. The Kolmogorov-Smirnov tests revealed no concerns in relation to normality and the Levene's Test indicated no problems with homogeneity of variance (F = 0.683, p = 0.638). No significant differences were revealed

between the attrition groups (F = 1.374, *df* = 2, 89, *p* = 0.258, r = 0.17) or the dropout groups (F = 1.170, *df* = 1, 89, *p* = 0.282, r = 0.11). The interaction was also non-significant (F = 0.658, *df* = 2, 89, *p* = 0.520, r = 0.12). The means are displayed in table 7.10 below.

Table 7.10: Pre-programme Aggressiveness by Attrition type (type one, type two, type three) and Group (non-starter, non-completer)

	Type one (N = 8)	Type two (N = 12)	Type three (N = 75)
Non-starters	4.17[a] (5.92[b]) [3[c]]	12.15 (8.00) [9]	16.13 (11.65) [43]
Non-completers	12.93 (16.01) [5]	16.67 (7.38) [3]	15.67 (12.29) [32]
Total	9.64 (13.30) [8]	13.28 (7.78) [12]	15.41 (11.02) [75]

[a] Mean [b] Standard Deviation [c] N

Previous Breach: Table 7.11 displays the distribution of previous breach by attrition groups. A chi-square analysis indicated that there was a no significant association between attrition group and previous breach (χ^2 (2, *N* = 167) = 5.383, *p* = 0.068, Φ_c = 0.18).

Table 7.11: Previous breach by attrition (type one, type two, type three) group

	No breach	Breach
Type one (%)	10 (47.6)	11 (52.4)
Type two (%)	9 (45.0)	11 (55.0)
Type three (%)	34 (27.0)	92 (73.0)
Total	53 (31.7)	114 (68.3)

Table 7.12 presents the distribution of previous breach by attrition and non-starter/non-completer groups. A chi-square analysis of the non-starter sample indicated that there was no significant association between attrition group and previous breach (χ^2 (2, *N* = 122) = 2.131, *p* = 0.344, Φ_c = 0.13). It was not possible to conduct a chi-square

analysis on the non-completer group as four cells (66.7%) had expected counts of less than five (Tabachnick & Fidell, 2006). Given that type three attrition represents 'true attrition' and the percentages of type one and type two offenders with previous breaches were so similar, the analysis was re-run with the type one and type two attrition groups collapsed into one category. This analysis revealed a significant association between attrition group and previous breach (χ^2 (2, N = 45) = 4.141, p = 0.042, Φ_c = 0.30). When evaluating this result, however, it should be borne in mind that the analysis returned one cell with an expected count of less than five. This amounted to 25% of the cells and hence the analysis should be considered unreliable. In the absence of further possibilities (i.e. further collapsing of categories) and due to only one cell returning expected counts of less than five, it was decided to present this information as indicative of a relationship.

Table 7.12: Previous breach by attrition (type one, type two, type three) and dropout group (non-starter, non-completer)

		No breach	Breach
Non-starters	Type one (%)	5 (38.5)	8 (61.5)
	Type two (%)	7 (41.2)	10 (58.8)
	Type three (%)	24 (26.1)	68 (73.9)
	Total	36 (29.5)	86 (70.5)
Non-completers	Type one (%)	5 (62.5)	3 (37.5)
	Type two (%)	2 (66.7)	1 (33.3)
	Type three (%)	10 (29.4)	24 (70.6)
	Total	17 (37.8)	28 (62.2)

Number of Previous Convictions (total and prior to age 18): Table 7.13 presents the mean number of previous convictions (total and prior to age 18) of the three attrition groups in total and by group (non-starter, non-completer). As with the age and OGRS2 variables, Kolmogorov-Smirnov tests revealed that both of these variable were non-normally distributed amongst one of the attrition groups (type three total convictions, *K-S Z* = 2.183, $p < 0.001$, and type three convictions aged under 18, *K-S Z* =

1.711, $p = 0.006$) and hence non-parametric tests were again used to test for differences between the attrition groups. A Kruskal-Wallis test revealed there to be no significant differences between the three groups' total number of previous convictions using the overall sample (χ^2 (2, $N = 167$) = 0.081, $p = 0.960$, $\Phi_c = 0.02$) or the number of convictions prior to the age of 18 (χ^2 (2, $N = 167$) = 1.799, $p = 0.407$, $\Phi_c = 0.10$).

Further Kruskal-Wallis tests on the separate groups of non-starters and non-completers revealed that there were no significant differences between the attrition groups within the non-starter sample in relation to total number of previous convictions (χ^2 (2, $N = 122$) = 0.063, $p = 0.969$, $\Phi_c = 0.02$) or in relation to the number of previous convictions prior to age 18 (χ^2 (2, $N = 122$) = 2.562, $p = 0.278$, $\Phi_c = 0.15$). There were also no significant differences between the attrition groups within the non-completer sample in relation to total number of previous convictions (χ^2 (2, $N = 45$) = 0.971, $p = 0.615$, $\Phi_c = 0.15$) or in relation to the number of previous convictions prior to age 18 (χ^2 (2, $N = 45$) = 0.564, $p = 0.754$, $\Phi_c = 0.11$).

Table 7.13: Total number of previous convictions and number of previous convictions prior to age 18 by attrition (type one, type two, type three) and dropout group (non-starter, non-completer)

		Type one (N = 21)	Type two (N = 20)	Type three (N = 126)
Non-starters	Total	9.46[a] (7.22[b]) [13[c]]	8.47 (5.63) [17]	9.21 (7.37) [92]
	< 18 years	4.92 (5.42) [13]	2.00 (2.06) [17]	3.12 (3.45) [92]
Non-completers	Total	9.50 (14.31) [8]	5.00 (5.20) [3]	7.97 (6.90) [34]
	< 18 years	2.63 (3.42) [8]	1.00 (1.00) [3]	3.21 (4.18) [34]
Total	Total	9.48 (10.15) [21]	7.95 (5.58) [20]	8.87 (7.24) [126]
	< 18 years	4.05 (4.80) [21]	1.85 (1.95) [20]	3.14 (3.65) [126]

[a] Mean [b] Standard Deviation [c] N

Age at First Conviction: Table 7.14 presents the mean age at first conviction of the three attrition groups in total and by group (non-starter, non-completer). Kolmogorov-Smirnov tests revealed that this variable was non-normally distributed

amongst one of the attrition groups (type three, *K-S Z* = 1.714, *p* = 0.006) and hence non-parametric tests were used to test for differences. The Kruskal-Wallis test revealed no significant differences between the three groups using the overall sample (χ^2 (2, *N* = 167) = 0.024, *p* = 0.988, Φ_c = 0.01). Further Kruskal-Wallis tests on the separate groups of non-starters and non-completers revealed there to be no significant differences between the attrition groups within the non-starter sample (χ^2 (2, *N* = 122) = 0.635, *p* = 0.728, Φ_c = 0.07) or within the non-completer sample (χ^2 (2, *N* = 45) = 1.188, *p* = 0.552, Φ_c = 0.16).

Table 7.14: Age at first conviction by attrition (type one, type two, type three) and dropout group (non-starter, non-completer)

	Type one (N = 21)	Type two (N = 20)	Type three (N = 126)
Non-starters	16.15[a] (3.81[b]) [13[c]]	16.71 (4.01) [17]	16.52 (2.90) [92]
Non-completers	17.38 (2.93) [8]	18.00 (2.65) [3]	16.56 (1.96) [34]
Total	16.62 (3.49) [21]	16.90 (3.81) [20]	16.53 (2.67) [126]

[a] Mean [b] Standard Deviation [c] N

Copas Rate: A three (type one, type two, type three) by two (non-starter, non-completer) between subjects ANOVA was conducted in relation to the Copas rate variable. The Kolmogorov-Smirnov tests revealed no concerns in relation to normality and the Levene's Test indicated no problems with homogeneity of variance (F = 1.881, *p* = 0.100). No significant differences were revealed between the attrition groups (F = 2.282, *df* = 2, 161, *p* = 0.105, r = 0.17) or the dropout groups (F = 0.728, *df* = 1, 161, *p* = 0.395, r = 0.07). The interaction was also non-significant (F = 0.189, *df* = 2, 161, *p* = 0.828, r = 0.04). The means are displayed in table 7.15 below.

Table 7.15: Copas rate by Attrition type (type one, type two, type three) and Group (non-starter, non-completer)

	Type one (N = 21)	Type two (N = 20)	Type three (N = 126)	Total (N = 167)
Non-starters	62.67[a] (15.77[b]) [13[c]]	50.20 (12.04) [17]	61.56 (19.57) [92]	60.09 (18.65) [55]
Non-completers	55.99 (19.05) [8]	45.75 (9.80) [3]	60.19 (18.68) [34]	58.48 (18.39) [40]
Total	60.12 (16.95) [21]	49.53 (11.61) [20]	61.19 (19.27) [122]	59.66 (18.54) [167]

[a] Mean [b] Standard Deviation [c] N

OASys Total Score: A three (type one, type two, type three) by two (non-starter, non-completer) between subjects ANOVA was conducted in relation to the OASys total score. The Kolmogorov-Smirnov tests revealed no concerns in relation to normality and the Levene's Test indicated no problems with homogeneity of variance ($F = 1.404$, $p = 0.226$). No significant differences were revealed between the attrition groups ($F = 0.441$, $df = 2, 161$, $p = 0.644$, $r = 0.07$) and the interaction between attrition and dropout groups was also non-significant ($F = 0.249$, $df = 2, 161$, $p = 0.780$, $r = 0.05$). There was a significant difference between the dropout groups, however ($F = 5.147$, $df = 1, 161$, $p = 0.025$, $r = 0.18$). The non-starters had significantly higher OASys total scores than the non-completers. The means are displayed in table 7.16 below.

Table 7.16: OASys total score by Attrition type (type one, type two, type three) and Group (non-starter, non-completer)

	Type one (N = 21)	Type two (N = 20)	Type three (N = 126)	Total (N = 167)
Non-starters	82.38[a] (34.73[b]) [13[c]]	75.88 (28.85) [17]	79.48 (25.27) [92]	79.29 (26.70) [55]
Non-completers	64.50 (31.11) [8]	55.33 (41.89) [3]	68.59 (26.84) [34]	66.98 (28.04) [40]
Total	75.57 (33.79) [21]	72.80 (30.69) [20]	76.54 (26.05) [122]	75.97 (27.53) [167]

[a] Mean [b] Standard Deviation [c] N

OASys risk classification: A chi-square analysis was undertaken on the full dataset. Table 7.17 presents the distribution of OASys risk classification by attrition group. It was not possible to conduct a chi-square analyses on the non-starter or non-completer groups as more than 20% of cells (44.4%) within each analysis had expected counts of less than five (Tabachnick & Fidell, 2006). As it did not make theoretical sense to collapse the categories, the analysis relating to these data was halted. The table below provides information relating to the distributions. However, again there were problems relating to the number of cells with counts less than five (44.4%). Again, as it made no theoretical sense to collapse the categories, the analysis was halted.

Table 7.17: OASys risk classification by attrition (type one, type two, type three) group

	Low	Medium	High
Type one (%)	3 (14.3)	11 (52.4)	7 (33.3)
Type two (%)	3 (15.0)	12 (60.0)	5 (25.0)
Type three (%)	8 (6.3)	95 (75.4)	23 (18.3)
Total	14 (8.4)	118 (70.7)	35 (21.0)

Table 7.18 displays the distribution of OASys risk classification by attrition and dropout groups. Chi-square analyses were attempted on the separate non-starter and non-completer samples. However, again there were problems relating to the number of cells with counts less than five (44.4% and 55.6% respectively). Again, as it made no theoretical sense to collapse the categories, the analysis was halted.

Table 7.18: OASys risk classification by attrition (type one, type two, type three) and dropout group (non-starter, non-completer)

		Low	Medium	High
Non-starters	Type one (%)	1 (7.7)	6 (46.2)	6 (46.2)
	Type two (%)	2 (11.8)	11 (64.7)	4 (23.5)
	Type three (%)	4 (4.3)	70 (76.1)	18 (19.6)
	Total	7 (5.7)	87 (71.3)	28 (23.0)
Non-completers	Type one (%)	2 (25.0)	5 (62.5)	1 (12.5)
	Type two (%)	1 (33.3)	1 (33.3)	1 (33.3)
	Type three (%)	4 (11.8)	25 (73.5)	5 (14.7)
	Total	7 (15.6)	31 (68.9)	7 (15.6)

OASys Criminogenic Need Score: A three (type one, type two, type three) by two (non-starter, non-completer) between subjects ANOVA was conducted in relation to the OASys criminogenic need score. As with the OASys total score, the Kolmogorov-Smirnov tests revealed no concerns in relation to normality and the Levene's Test indicated no problems with homogeneity of variance ($F = 1.260$, $p = 0.284$). No significant differences were revealed between the attrition groups ($F = 0.854$, $df = 2, 161$, $p = 0.427$, $r = 0.10$) and the interaction between attrition and dropout groups was also non-significant ($F = 0.284$, $df = 2, 161$, $p = 0.753$, $r = 0.06$). There was a significant difference between the dropout groups, however ($F = 4.360$, $df = 1, 161$, $p = 0.038$, $r = 0.16$). The non-starters had significantly higher OASys criminogenic need scores than the non-completers. The means are displayed in table 7.19 below.

Table 7.19: OASys criminogenic need score by Attrition type (type one, type two, type three) and Group (non-starter, non-completer)

	Type one (N = 21)	Type two (N = 20)	Type three (N = 126)	Total (N = 167)
Non-starters	40.85[a] (24.34[b]) [13[c]]	34.06 (22.54) [17]	36.39 (21.18) [92]	36.54 (21.59) [55]
Non-completers	30.88 (17.52) [8]	17.00 (13.08) [3]	29.21 (17.05) [34]	28.69 (16.88) [40]
Total	37.05 (22.08) [21]	31.50 (22.02) [20]	34.45 (20.34) [122]	34.43 (20.68) [167]

[a] Mean [b] Standard Deviation [c] N

Number of Attempts at ETS on this Order: Table 7.20 presents the mean number of attempts at ETS of the three attrition groups in total and by group (non-starter, non-completer). Kolmogorov-Smirnov tests revealed that this variable was non-normally distributed amongst two of the attrition groups (type two, *K-S Z* = 1.540, *p* = 0.017; type three, *K-S Z* = 2.695, *p* < 0.001) and hence non-parametric tests were used to test for differences between the attrition groups. The Kruskal-Wallis test revealed no significant differences between the three groups using the overall sample (χ^2 (2, *N* = 167) = 3.189, *p* = 0.203, Φ_c = 0.14). Further Kruskal-Wallis tests on the separate groups of non-starters and non-completers revealed there to be no significant differences between the attrition groups within the non-starter sample (χ^2 (2, *N* = 122) = 2.825, *p* = 0.244, Φ_c = 0.15) or within the non-completer sample (χ^2 (2, *N* = 45) = 3.677, *p* = 0.159, Φ_c = 0.29).

Table 7.20: Mean number of attempts at ETS on this order by Attrition type (type one, type two, type three) and Group (non-starter, non-completer)

	Type one (N = 21)	Type two (N = 20)	Type three (N = 126)
Non-starters	0.23[a] (0.44[b]) [13[c]]	0.41 (0.80) [17]	0.58 (0.82) [92]
Non-completers	1.88 (0.64) [8]	1.00 (0.00) [3]	1.79 (0.95) [34]
Total	0.86 (0.96) [21]	0.50 (0.76) [20]	0.90 (1.01) [122]

[a] Mean [b] Standard Deviation [c] N

OGRS2 Appropriateness: Table 7.21 presents the distribution of OGRS2 appropriateness by attrition group. A chi-square analysis was undertaken on the full dataset. However, again there were problems relating to the number of cells with counts less than five (33.33%)[24]. As such, it was decided to collapse the appropriateness categories into 'appropriate' and 'not appropriate'. The too low and too high categories were hence merged. There was a no significant association between attrition group and OGRS2 appropriateness (χ^2 (2, N = 167) = 4.342, p = 0.114, Φ_c = 0.16).

Table 7.21: OGRS2 appropriateness by attrition (type one, type two, type three) group

	Too low	Appropriate	Too high
Type one (%)	1 (4.8)	11 (52.4)	9 (42.9)
Type two (%)	2 (10.0)	15 (75.0)	3 (15.0)
Type three (%)	1 (0.8)	63 (50.0)	62 (49.2)
Total	4 (2.4)	89 (53.3)	74 (44.3)

Table 7.22 displays the distribution of OGRS2 appropriateness by attrition and dropout groups. It was not possible to conduct a chi-square analyses on the non-starter or non-completer groups as more than 20% of cells (33.3%[25] and 77.87% respectively)

[24] Although unreliable this chi-square was significant (χ^2 (2, N = 167) = 13.267, p = 0.01, Φ_c = 0.28)
[25] Although unreliable the chi-square on the non-starter sample was significant (χ^2 (2, N = 122) = 14.343, p = 0.006, Φ_c = 0.34)

within each analysis had expected counts of less than five (Tabachnick & Fidell, 2006). As such, the categories were collapsed as above and a further chi-square was undertaken with the recoded 'appropriate' and 'not appropriate' categories. The chi-square relating to the non-starters was non-significant (χ^2 (2, *N* = 122) = 2.445, *p* = 0.294, Φ_c = 0.14). However, the chi-square relating to the non-completers was still unreliable with 66.67% of cells still with counts less than five. As it did not make theoretical sense to collapse the categories any further, the analysis relating to these data was halted.

Table 7.22: OGRS2 appropriateness by attrition (type one, type two, type three) and dropout group (non-starter, non-completer)

		Too low	Appropriate	Too high
Non-starters	Type one (%)	1 (7.7)	7 (53.8)	5 (38.5)
	Type two (%)	2 (11.8)	12 (70.6)	3 (17.6)
	Type three (%)	0 (0.0)	46 (50.0)	46 (50.0)
	Total	3 (2.5)	65 (53.3)	54 (44.3)
Non-completers	Type one (%)	0 (0.0)	4 (50.0)	4 (50.0)
	Type two (%)	0 (0.0)	3 (100.0)	0 (0.0)
	Type three (%)	1 (2.9)	17 (50.0)	16 (47.1)
	Total	1 (2.2)	24 (53.3)	20 (44.4)

Correlates of 3A, 3B, and 3C attrition

Given that the type three attrition group was the largest of the three and it consisted of three types of attrition (3A, 3B, and 3C), it was decided to run analyses to determine whether there were any differences between these three groups in relation to the above variables. Given the differences observed both within this and previous chapters between the non-starters and non-completers it was decided not to merge the samples into one group. However, this meant that the following analysis could only be undertaken with the non-starter sample as the non-completer group sample size was too small for meaningful analysis (3A = 6, 3B = 6, and 3C = 22).

Age: Table 7.23 presents the mean ages of the three groups. Kolmogorov-Smirnov tests revealed that this variable was non-normally distributed within one of the attrition groups (3B, *K-S Z* = 1.513, *p* = 0.020) and hence non-parametric Kruskal-Wallis tests were conducted to test for differences in age between the attrition groups. There were no significant differences between the groups (χ^2 (2, *N* = 92) = 5.890, *p* = 0.053, Φ_c = 0.25).

Table 7.23: Age and OGRS2 by Attrition type (3A, 3B, 3C) amongst the non-starter group

	3A (N = 15)	3B (N = 52)	3C (N = 25)
Age	25.60[a] (7.22[b])	22.42 (5.48)	24.20 (4.44)
OGRS2	61.60 (21.84)	72.56 (18.17)	74.92 (20.05)

[a] Mean [b] Standard Deviation

OGRS2: Table 7.23 also presents the mean OGRS2 scores of the three groups. There were no problems with normality, as assessed using the Kolmogorov-Smirnov test, or with the homogeneity of variance assumption (Levene's statistic = 1.112, *p* = 0.333) As such a one-way ANOVA was undertaken. No significant differences were observed between the groups: F = 2.443, *df* = 2, 89, *p* = 0.093, r = 0.23.

Pre-programme Racial Intimacy: A one-way ANOVA was conducted to test for differences between the groups in relation to the pre-programme racial intimacy variable. The Kolmogorov-Smirnov tests revealed no problems with normality and the Levene's Test indicated no problems with homogeneity of variance (F = 0.714, *p* = 0.496). No significant differences were revealed between the groups (F = 0.396, *df* = 2, 38, *p* = 0.676, r = 0.14). The means are displayed in table 7.24 below.

Table 7.24: Pre-programme Racial Intimacy and Aggressiveness by Attrition type (3A, 3B, 3C) amongst the non-starter group

	3A	3B	3C
Racial Intimacy	15.00[a] (3.87[b]) [11[c]]	14.56 (3.96) [18]	16.00 (5.33) [12]
Aggressiveness	15.50 (15.77) [12]	16.33 (10.47) [20]	16.48 (9.33) [11]

[a] Mean [b] Standard Deviation [c] N

Pre-programme Aggressiveness: A one-way ANOVA was conducted to test for differences between the groups in relation to the pre-programme aggressiveness variable. The Kolmogorov-Smirnov tests revealed no problems with normality and the Levene's Test indicated no problems with homogeneity of variance (F = 0.822, *p* = 0.447). No significant differences were revealed between the groups (F = 0.024, *df* = 2, 40, *p* = 0.976, r = 0.03). The means are displayed in table 7.24 above.

Previous Breach: Table 7.25 displays the distribution of previous breach by attrition groups. A chi-square analysis indicated that there was a significant association between attrition group and previous breach (χ^2 (2, *N* = 92) = 10.036, *p* = 0.007, Φ_c = 0.33). The incidence of previous breach was higher in the 3B and 3C categories than in the 3A category.

Table 7.25: Previous breach by attrition group (3A, 3B, 3C) amongst the non-starter group

	No breach	Breach
3A (%)	8 (53.3)	7 (46.7)
3B (%)	14 (26.9)	38 (73.1)
3C (%)	2 (8.0)	23 (92.0)
Total	24 (26.1)	68 (73.0)

Number of Previous Convictions (total and prior to age 18): Table 7.26 presents the mean number of previous convictions (total and prior to age 18) of the three attrition groups. The Kolmogorov-Smirnov tests revealed no problems with normality. However, the Levene's Test indicated that the assumption of homogeneity of variance was violated (F = 6.422, *p* = 0.002) and hence non-parametric tests were used in relation to this variable. A Kruskal-Wallis test revealed there to be no significant differences between the three groups' (χ^2 (2, *N* = 92) = 3.168, *p* = 0.205, Φ_c = 0.19).

A one-way ANOVA was conducted to test for differences between the groups in relation to the total number of previous convictions variable. The Kolmogorov-Smirnov tests revealed no problems with normality and the Levene's Test indicated no problems with homogeneity of variance (F = 1.250, *p* = 0.291). Significant differences were observed between the groups (F = 3.122, *df* = 2, 89, *p* = 0.049, r = 0.26). Despite this overall significant difference, Scheffe post-hoc tests did not reveal any significant pairwise differences between the attrition groups; the difference between groups 3B and 3C approached significance (*p* = 0.062) whilst the other comparisons were far from significant (3A and 3B: *p* = 0.999; 3A and 3C: *p* = 0.197).

Table 7.26: Total number of previous convictions and number of previous convictions prior to age 18 by attrition group (3A, 3B, 3C) amongst the non-starter group

	3A (N = 15)	3B (N = 52)	3C (N = 25)
Total	8.00[a] (8.20[b])	8.08 (6.85)	12.28 (7.31)
< 18 years	1.53 (1.64)	3.15 (3.10)	4.00 (4.57)

[a] Mean [b] Standard Deviation

Age at First Conviction: A one-way ANOVA was conducted to test for differences between the groups in relation to the age at first conviction variable. The Kolmogorov-Smirnov tests revealed no problems with normality and the Levene's Test indicated no problems with homogeneity of variance (F = 0.625, *p* = 0.538). A significant difference was revealed between the groups (F = 4.267, *df* = 2, 89, *p* = 0.017, r = 0.29). Post-hoc Scheffe tests indicated significant differences between the 3A and 3C attrition groups (*p*

= 0.017). The 3A group were significantly older at their first conviction than the 3C group. The means are displayed in table 7.27 below.

Table 7.27: Age at first conviction, Copas rate, OASys total score and OASys criminogenic need score by attrition group (3A, 3B, 3C) amongst the non-starter group

	3A (N = 15)	3B (N = 52)	3C (N = 25)
Age at first conviction	18.27[a] (3.45[b])	16.46 (2.89)	15.60 (2.12)
Copas rate	53.20 (19.72)	59.96 (17.57)	69.89 (21.15)
OASys total score	64.00 (20.25)	82.42 (25.54)	82.64 (24.85)
OASys criminogenic score	25.87 (13.92)	41.21 (22.59)	32.39 (21.18)

[a] Mean [b] Standard Deviation

Copas Rate: A one-way ANOVA was conducted to test for differences between the groups in relation to their Copas rate. Kolmogorov-Smirnov tests revealed no problems with normality and the Levene's Test indicated no problems with homogeneity of variance (F = 1.270, *p* = 0.286). Significant differences were observed between the groups (F = 4.067, *df* = 2, 89, *p* = 0.02, r = 0.29). Post-hoc Scheffe tests indicated significant differences between the 3A and 3C attrition groups. The 3A group had a significantly lower Copas rate than the 3C group. The means are displayed in table 7.27 above.

OASys Total Score: A one-way ANOVA was conducted to test for differences between the groups in relation to their OASys total score. Once again Kolmogorov-Smirnov tests revealed no problems with normality and the Levene's Test indicated no problems with homogeneity of variance (F = 0.429, *p* = 0.653). A significant difference as revealed between the groups (F = 3.550, *df* = 2, 89, *p* = 0.033, r = 0.27). Post-hoc Scheffe tests indicated significant differences between the 3A and 3B attrition groups. The 3B group had significantly higher OASys total scores than the 3A group. The means are displayed in table 7.27 above.

OASys risk classification: Table 7.28 displays the distribution of OASys risk classifcations by attrition groups. A chi-square analysis indicated that 55.6% of cells had expected of counts less than five. As such, the chi-square was deemed unreliable (Tabachnick & Fidell, 2006). As it did not make theoretical sense to collapse the categories, the analysis was halted at this point.

Table 7.28: OASys risk classification by attrition group (3A, 3B, 3C) amongst the non-starter group

	Low	Medium	High
3A (%)	3 (20.0)	12 (80.0)	0 (0.0)
3B (%)	1 (1.9)	39 (75.0)	12 (23.1)
3C (%)	0 (0)	19 (76.0)	6 (24.0)
Total	4 (4.3)	70 (6.1)	18 (19.6)

OASys Criminogenic Need Score: Kolmogorov-Smirnov tests revealed that the OASys criminogenic need scores were normally distributed and the Levene's test indicated no issues with homogeneity of variance (F = 2.407, p = 0.096). A one-way ANOVA was therefore conducted to test for differences between the groups. A significant difference as revealed between the groups (F = 3.804, df = 2, 89, p = 0.026, r = 0.28). Post-hoc Scheffe test revealed significant differences between the 3A and 3B groups (p = 0.044): the 3B group had significantly more criminogenic needs than the 3A group. There were no other significant pairwise differences (3A and 3C, p = 0.599; 3B and 3C, p = 0.239). The means are displayed in table 7.27 above.

Number of Attempts at ETS on this Order: Table 7.29 presents the mean number of attempts at ETS of the three attrition groups. Kolmogorov-Smirnov tests revealed that this variable was non-normally distributed amongst two of the attrition groups (3B, *K-S Z* = 2.743, $p < 0.001$; 3C, *K-S Z* = 1.341, p = 0.055) and hence non-parametric tests were used to test for differences between the attrition groups. The Kruskal-Wallis test revealed no significant differences between the three groups (χ^2 (2, N = 92) = 2.674, p = 0.263, Φ_c = 0.17).

Table 7.29: Mean number of attempts at ETS on this order by attrition group (3A, 3B, 3C) amongst the non-starter group

	3A (N = 15)	3B (N = 52)	3C (N = 25)
Number of attempts	0.67[a] (0.90[b])	0.44 (0.67)	0.80 (1.00)

[a] Mean [b] Standard Deviation

OGRS2 Appropriateness: Table 7.30 presents the distribution of OGRS2 appropriateness by attrition group. There were no cases that fell into the too low group and hence this category was removed. A chi-square analysis was undertaken on the full dataset. There was a no significant association between attrition group and previous breach (χ^2 (2, *N* = 92) = 0.960, *p* = 0.619, Φ_c = 0.10).

Table 7.30: OGRS2 appropriateness by attrition group (3A, 3B, 3C) amongst the non-starter group

	Appropriate	Too high
3A (%)	9 (60.0)	6 (40.0)
3B (%)	26 (50.0)	26 (50.0)
3C (%)	11 (44.0)	14 (56.0)
Total	46 (50.0)	46 (50.0)

Discussion

The purpose of this chapter was twofold: first, to classify programme dropouts according to the action-orientated typology of Stephens (2003) and Stephens and Turner (2004), and second, to determine whether these types of programme dropouts differed from one another in relation to demographic, psychometric, need, and criminogenic factors. The first of these aims would provide information on the reasons for attrition and whether these related to factors internal to the offender or within the organisation. The second aim would provide indications as to whether there are further subsets of offenders within the non-starters and non-completers. This information would inform not only those working with these offenders but also the research community. If differences were found these would need to be considered in future attrition research.

The findings presented in this chapter attest to the findings of Stephens and Turner (2004) in that the vast majority of attrition was classified as type three: these offenders could, should, but didn't attend the programme to which they were sentenced. In taking the coding of type three attrition further, it was also possible to determine that for ten percent of all those sentenced to attend ETS (or seven percent when using the final completion status), a group placement that met the needs of these offenders was not available , either due to clashes with employment or child care arrangements or due to probation administrative mistakes. When these figures are considered alongside the type one (couldn't attend) and type two (shouldn't attend) attrition, it is seen that over a third of all programme dropouts either couldn't, shouldn't, or were not able due to the unavailability of a programme place, to attend the ETS programme. The remainder of attrition was due to either a breach of appointments other than those relating to the ETS programme (34.73% of dropouts) or related to those who were apparently compliant with other aspects of their order but dropped out of the ETS programme for other reasons, such as illness, lateness, breach of the behavioural requirements, or general non-attendance (27.81% of dropouts).

The proportions of those who could not (7.12%) or should not (6.73%) attend the intervention were similar to those reported by Stephens and Turner (2004): within their

research on the ETS programme these figures were seven and five percent respectively. Half of the offenders who could not attend were in custody, a quarter had moved out of the area, and the remainder were not able to attend due to a court or breach appearance. Whilst such information does provide indications as to why a sentenced offender did not attend the ETS programme, it does not, however, provide the full picture: had those offenders in custody been detained for a further offence or for a breach of their current order? If it was the latter, did this relate to their non-attendance at the ETS programme? Likewise, do the breach/court proceedings relate to non-attendance at the ETS programme or some other incident? From the case notes, it was not possible to determine the answers to these questions but in order to fully unpick the reasons for attrition, such information is needed. As such, it is perhaps too early to conclude, as Stephens and Turner did, that "It is likely that no further action would have been possible in these cases" (p. 22).

More firm conclusions can, however, be drawn from the information relating to the reasons for type two attrition. Those that could but should not attend the ETS programme were either assessed at post-sentence stage as being unsuitable for the programme or had been subject to administrative mistakes. Such findings give weight to the arguments of Turner (2006) and those contained within the Sussex Probation Area (2003) report in relation to the need for service-wide training relating to the appropriate targeting of offenders for accredited programmes. The very fact that further assessment determined that these offenders were unsuitable for the ETS programme indicates that such inappropriate referrals are preventable through careful assessment and consultation. Appropriate training should therefore focus on ensuring such referrals are prevented prior to sentencing. Given the cautious suggestions within the literature (Hollin et al., 2008; McMurran & Theodosi, 2007; and within chapter two above) that the process of programme dropout may be detrimental to reconviction outcomes, such referrals should be minimised.

A comparison of the non-starters and non-completers shows that the non-completers were more likely than non-starters to be type one dropouts (17.78 vs 10.66) and less likely to be type two dropouts (0.07 vs 13.93). This latter finding indicates that the further assessment which determines unsuitability is, in most cases, carried out prior to commencement on the programme. In relation to type three attrition, the

proportions of non-completers and non-starter for whom a programme place was not available were very similar at approximately one in eight. The vast majority of these were not able to continue on their programme due to conflicts with employment. Whilst the probation area within which the evaluation was undertaken does provide programmes at different times of the day to accommodate for employment patterns, it would seem that the conclusions of Stephens and Turner still (2004) resonate: greater flexibility is required in order to reduce the attrition rate.

Non-starters were more likely than non-completers to have dropped out of ETS due to being in breach of other appointments: almost half of all those who failed to commence did so due to breach of other appointments. It would seem therefore that non-starters were generally either unavailable, screened out, or did not commence the ETS programme due to breach of other appointments: only one in five were available, suitable and apparently compliant with their order but then failed to attend ETS. In contrast, half of the non-completers were in a position to attend the ETS programme and were compliant with their order but did not complete the programme. Of these half failed to attend the required number of sessions whilst a quarter were removed from the programme due to behavioural issues. The difference between the non-starters and non-completers thus seems to lie in the point at which they withdraw their engagement with the probation process. Whilst the majority of non-starters withdraw prior to the programme, perhaps meaning that they never reach the position where they are offered a programme place, the non-completers seem to be compliant with their order up to the point of the programme. A qualitative research programme would perhaps provide more evidence in relation to this finding but from the data presented here it would seem that non-starters fail to engage with probation whilst the non-completers fail to engage with the programme. Such lack of programme engagement amongst the non-completers could indicate issues with responsivity: further research within this area would be welcomed.

In relation to the correlates of the different types of attrition, type one and three non-starters were significantly younger and had higher risk of reconviction scores (perhaps due, in part at least, to the fact that they were younger) than the type two non-starters. At first consideration this finding is perhaps not surprising: type two non-starters are only such due to their inappropriate referral to the programme whilst type

one and type three dropouts are more likely to be in breach, in custody, or chose not to attend the programme and, as such, perhaps represent true non-starters, a group which the literature consistently shows to be more criminogenic in nature than those who comply with their programme order (Craissati & Beech, 2001; Nunes & Cortonni, 2006a; Turner, 2006; Wormith & Olver, 2002). This finding, however, only held for the non-starters and not for the non-completers. Closer consideration of the data indicates potential reasons for this disparity. First, the non-completer sample size is much smaller than that of the non-starters (45 and 122 respectively). As such, it is possible that a lack of power resulted in a lack of statistical significance. Indeed the pattern observed amongst the non-starters in relation to risk scores is mirrored in the data of the non-completers: the type two group have lower scores than type one and three groups. Furthermore, the effect sizes indicate the potential existence of such differences: there is little disparity between the effect size for the non-starters ($\Phi_c = 0.28$) and the non-completers ($\Phi_c = 0.23$). However, the same does not follow in relation to the age variable (non-starters, $\Phi_c = 0.30$ and non-completers, $\Phi_c = 0.05$). This might indicate that the differences in OGRS2 scores amongst the non-starters may be due, in part, to the differences in ages of the groups. However, age is but one contributor to the OGRS2 score algorithm and it would appear that amongst the non-completers it is more likely that other factors influence the differences in risk of reconviction scores between the groups. Further research is needed to unpick this issue further.

In relation to compliance amongst the sample, there was an association between whether the offender had previously been breached and the type of attrition that they now presented. Two thirds of type three dropouts, compared with just a third of type one and type two offenders, had previous problems with order compliance. Such a pattern was observed across the sample and amongst the separate groups of non-completers and non-starters (although for this group the difference was less pronounced due to increased compliance issues across all attrition types rather then reduced attrition within the type three group). Analysis of the non-starter 3A, 3B, and 3C groups indicated that this problem was more prevalent amongst the latter two groups: those offenders who failed to commence because there was no group place available to them had a lower incidence of breach comparative to those who were either non-compliant prior to or with the programme. Programmes teams should

perhaps bear such previous behaviour in mind when working with their caseload: of those dropouts who could and should attend their interventions dropout, three in four have previously been subject to breach proceedings of some kind. Comparison with the completers figures in chapter five (p. 145), however, shows that breach also features in the history of over half of the programme's completers. As such, the presence of previous breach is not an indicator of dropout per se; further research could usefully investigate this issue to determine whether programme dropouts (and programme dropouts of different types) have a greater incidence or are more likely to have experienced a particular type of breach than the programme completers.

As such, it would seem that sub-groups of programme dropouts exist within programme non-starters and non-completers. First, there is an organisational component to attrition: some offenders are wrongly assessed as suitable for the programme (type two) or they do not have a place available to them (type 3A). Second, there are indications[26] that those who do have a place available to them but still fail to complete (3B and 3C) are younger, more criminogenic, more likely to have a history of non-compliance, commenced offending at an earlier age, offended at a faster pace, and had more criminogenic needs than the other types of dropouts. Previous chapters have concluded that both non-completers and non-starters are more criminogenic than the programme completers. Within this chapter, there have been indications of the existence of further subgroups of non-starters and non-completers and that these can be differentiated by the reasons for their dropout. The finding that these groups also differ in respect of their 'criminogenic-ness' indicates that those offenders representing 'true attrition' (i.e. those who could and should attend, for whom a place was available but still did not attend) are the most criminogenic of all. Such findings should be borne in mind when evaluating offending behaviour programmes, their effectiveness and the impact of programme dropout. Such investigations should consider why an offender was not able to attend the programme to which they were referred and whether the programme outcomes between these attrition groups (after controlling or matching for criminogenic variables) differ.

[26] Although some of these were non-significant, the means indicated such patterns and the effect sizes indicated a potential lack of power.

There are a number of limitations to bear in mind when evaluating the results of the research presented within this chapter and the implications stemming from them. The findings of applied research such as this can only ever be as good as the data upon which it is based. The data presented here were taken from probation databases which were developed for the monitoring of offenders' progress through and compliance with their sentence rather than for evaluation. As such, the information held within them is not always within the format required for such research and hence the use of such data can undermine the research aims. In addition and as argued by Stephens (2003), it is likely that attrition is not the outcome of one sole reason. In reality, there are probably a multitude of factors that result in this outcome. As such there are elements of subjectivity that must be accepted when conducting such research, first in the recording of reasons for dropout by probation staff, and second, in the classification of those reasons into Stephen's typology by the researcher. Stephens herself does argue that some reasons could easily and validly be classified within two or more categories. An example of this is the coding of 'in court/breach appearance' as a type one reason for non-attendance on the ETS programme. The decision was made for the purposes of this thesis to code this reason as a type one reason for attrition as it was not physically possible for the offender to attend the programme. However, it would perhaps have been just as valid to code this as a type 3B reason: in being subject to breach proceedings or in court the offender had (potentially) failed to comply with other aspects of their order. This element of subjectivity should therefore be borne in mind when evaluating the results of this analysis.

This research has also suffered from a lack of power. Although the original dataset of almost 300 offenders is of decent proportions, the coding into groups and sub-groups inevitably results in smaller and smaller samples sizes within the subgroups. As such, some analyses were either not possible (non-completer 3A, 3B, and 3C correlates) or were invalid (some of the chi-squares). To counter this, however, effect sizes were calculated throughout (and indeed the disparity between some of the returned p values and the effect sizes attest to the lack of power) and the results were discussed in respect of these where appropriate.

Chapter Eight

General Discussion

General Discussion

This thesis aimed to investigate the issue of attrition from community based general offending behaviour programmes so as to inform both the research knowledge base and clinical practice. That attrition from community based programmes is a substantial issue has been well established within the literature (Hollin et al., 2004, 2008; Hollis, 2007; Kemshall & Canton, 2002; Palmer et al., 2007, 2008, 2009). However, the impact of such attrition on reconviction outcomes has been a matter of some debate. Whilst a recent meta-analysis (McMurran & Theodosi, 2007) proffered the view that non-completion can be detrimental to outcomes, other commentators (e.g Debidin & Lovbakke, 2005) have argued that the observed outcomes are a production of self-selection and all programmes serve to do is to separate out those that 'would do well' (completers) from those that will be reconvicted (dropouts).

Initially this thesis aimed to contribute to the evidence base by evaluating the impact of attrition on reconviction outcomes using quasi-experimental methods. In the absence of randomisation, however, and recognising the need to evaluate the dropouts against a representative portion of the comparison group (Seager, Jellicoe & Dhaliwal, 2004), a matched analysis was undertaken. Using a national sample of offenders who were sentenced to one of two general offending behaviour programmes, matching was undertaken on a one to one basis to comparison group offenders who were subject to probation only. This methodological design was hence different to those seen within the majority of quasi-experimental research within this field. Previous research has, in the main, chosen to statistically control for differences between the experimental and comparison groups and when comparing outcomes has used the full comparison group as the yardstick for the naturally occurring groups of completers, non-completers, and non-starters (e.g. Hollin et al., 2004, 2008; McGuire et al., 2008; Palmer et al., 2007). The selection of the current methodology therefore had two advantages: first, it ensured that the comparison group and experimental group did not differ on key variables, and second, it allowed for the compartmentalising of the comparison group so that the naturally occurring sub-groups of completers, non-completers and non-starters could be compared with their one to one matches.

When considering the odds ratios produced by this analysis, there were tentative indications within these data of a completion effect: those who completed programmes were less likely to be reconvicted than the full comparison group. In addition, there were equally tentative suggestions of a non-completion effect: those who dropped out of programmes (both non-starters and non-completers) were more likely to be reconvicted than the full comparison group. Comparisons against the matched sections of the comparison group were underpowered[27] and hence failed to reach significance but again the odds ratios hinted at the conclusion that programme completers perform better than their matched comparisons and that programme non-completers perform worse than their one to one matches.

Given that the patterns in these matched analyses were similar to those observed within the wider literature (McMurran & Theodosi, 2007; Hollin et al., 2008; Palmer et al., 2007), chapter three explored the data for potential differences between the completers, non-completers, and non-starters of programmes. In the absence of randomisation, it is not possible to state with confidence that any differences in outcomes between the groups are due to the intervention (or the process of failing to complete the intervention). As such it becomes necessary to evaluate the differences between the groups so that subsequent research can match the experimental and comparison groups on, or statistically control for the influence of, these variables. The aim within this chapter therefore was to advance knowledge in relation to such differences but also to determine whether the patterns in outcomes as described above remained once any potential differences were statistically controlled for.

[27] It is assumed that the analyses are underpowered and that, as a consequence, the probability of a Type II error (failure to reject the null hypothesis when the null hypothesis is false) is elevated. However, power analyses have not been undertaken in relation to the samples in this thesis. The reason for this lies in the nature of the acquisition of the data. The data utilised in Part One of the thesis utilised a pre-existing sample and hence a priori power analyses were not possible. Similarly, the data used in Part Two comprised a population of offenders referred to a particular programme type within one probation area within a particular time period. As such a priori power analyses were not considered necessary as the sample size was already limited by the retrospective nature of the research; to know prior to data collection the sample size required would have been redundant as the size of the sample was limited. Additionally, whilst there was the opportunity to conduct post hoc power analyses, these are considered by some to be controversial due to their unreliability (e.g. Colegrave & Ruxton, 2003; Levine & Ensom, 2001; Thomas, 1997; Yuan & Maxwell, 2005). For example, Yuan and Maxwell reported that estimated power does not always correspond with actual power especially when the actual power is small and Thomas argues that different retrospective analyses can yield considerably different estimates of power. As such, retrospective or a priori power analyses were not undertaken within this thesis.

The univariate analyses within chapter three tested for differences between the groups in relation to demographic (including ratings of motivation and literacy problems), offender need (criminogenic and non-criminogenic), and offender programme suitability (or level of cognitive deficit) measures. The only variable that was found to differentiate the groups was the risk of reconviction, or OGRS2, score (Taylor, 1999). The chapter utilised this variable in two ways. First, as a purely continuous variable: completers had lower risk of reconviction scores than the non-completers and the non-starters. Second, as an indication of programme eligibility: the general offending behaviour programmes evidence-based targeting criteria state that medium to high risk offenders (defined as having OGRS2 scores of between 31 and 74) should be targeted for the programme (National Management Manual, 2000). In line with the findings of Palmer and colleagues (2008, 2009) nearly half of the sample was found to fall outside of the appropriate risk of reconviction banding for general offending behaviour programmes. There was also a significant association between appropriateness (too low, appropriate, too high) and programme attrition: of those who had scores above the targeting criteria half failed to commence the programme, within the appropriate banding the non-starter rate decreased to a third, and within the too low category it was just one in five.

Despite the arguments of Debidin and Lovbakke (2005) concerning the likely influence of motivation and criminogenic need on programme completion, dropout and subsequent outcomes, no differences were observed between the groups in relation to these variables. It should be noted, however, that the motivational assessments were utilised within this thesis were crude - probation officers were asked to provide ratings of their clients' motivation to attend the programme – and that the assessments of need (within chapter three at least) were undertaken using a perhaps underdeveloped precursor to the full risk and need assessment tool, OASys. As such, these conclusions should be viewed as tentative and further research should explore them in more detail.

The multivariate analyses within chapter three showed that the risk of reconviction score predicted programme completion (non-starter, non-completer, completer). This raised the possibility that any differences in reconviction outcomes between the three groups may be a function of risk of reconviction rather than a function of the programme. As such, this was tested within an additional logistic

regression analysis. After controlling for the risk of reconviction, the effect of group (completer, non-completer, non-starter) on reconviction outcomes remained. This indicates that whilst the completers, non-completers and non-starters do differ in relation to the risk of reconviction variable, this variable is not responsible for all of the variance in the outcomes of the groups. From the evidence to date, it is not possible to say that it is the impact of the programme (or dropout from the programme) which has caused the observed effects however; there is still the possibility that key unmeasured variable(s) that differ between the groups are also responsible for observed differences in reconviction outcomes. However, the elimination of variables within chapter three could perhaps be said to increase the chances that the observed outcomes are a function of the programme (and/or programme dropout).

The second part of this thesis set about to understand the completer, non-completer and non-starter groups in more detail; are there any differences between the groups that could inform the intervention outcome debate or the treatment readiness literature (Ward et al., 2004)? As such, over a two year period, offenders with sentences to attend the ETS programme within one probation area were studied. Individual (demographic, psychometric, offence-related, criminogenic and non-criminogenic need) and process factors (e.g. type of probation order, length of time between sentence and programme start, appropriateness of targeting, intensity and timing of programme) were considered to determine which, if any, related to attrition and its different forms. As such, this thesis aimed to add original knowledge to the research base by comparing programme completers with both programme non-completers *and* non-starters on a variety of variables relating to both the individual *and* process. Previous research has generally compared programme completers to non-completers or programme starters to non-starters and has tended to focus on either offender *or* process factors. As such, this thesis has presented a more rounded picture of attrition from offending behaviour programmes within community services than was previously available.

As seen elsewhere within the literature, programmes completers were found to be older and at a lower risk of reconviction than programme non-starters and non-completers (Craissati & Beech, 2001; Browne, Foreman, & Middleton, 1998; Hazeltine et al, 2002; Mosher & Phillips, Nunes & Cortoni, 2006a; Turner, 2006; van Voorhis et al,. 2004; Wormith & Olver, 2002; Zanis et al., 2003). Furthermore, completers were less

aggressive and had less potentially discriminatory attitudes than programme dropouts. As such, there were indications within these data that programme dropouts are more anti-social in nature and, potentially, in personality than programme completers.

In relation to the offence histories of the groups, the non-starters appeared to be more criminogenic than both the completers and the non-completers with the completers having the least criminogenic lifestyle. The analysis of the rate of offending between the groups, however, reveals that the non-completers had just as high a rate of offending as the non-starters, both of which are higher than that of the completers. Given that the non-completers are younger than both completers and non-starters, it is therefore possible that the non-completers are younger versions of the non-starters: they are just as criminogenic but have had less time to accumulate a history.

When considering the differences between those who commence a programme (completers and non-completers) and those that do not (non-starters), the non-starters were more likely to have had previous experiences of breach, had more previous convictions as a juvenile, commenced their criminality at a younger age, and had a higher rate of offending. As such, it would seem that the non-starters represent a type of offender that is distinct from the programme completer. Given the age differences between the groups, it could be hypothesised that the programme completers are those that have reached an age or maturity where they wish to desist from offending (e.g. Farrington, 1992, 1993; Moffit, 1993) and see the intervention as one way of helping them accomplish this. As such, the non-completers and non-starters could be seen as those who are yet to reach that age or maturity. However, given the differences in criminal history reported above, it would seem that this is not the case. Rather than being at a different point in the life course than the programme completers, it would seem that the non-starters (and possibly the non-completers) are instead a different type of offender altogether.

That is not to say, however, that all programme attrition is due to individual factors such as the ones discussed above. Chapters six and seven saw that process or organisational factors also play their part in the failure of an offender to attend or complete the programme to which they are sentenced. Chapter six was able to replicate the findings relating to inappropriate targeting which were unearthed in chapter three: over a third of the sample had OGRS2 scores above the targeting upper limit of 74 and

seven in ten of these failed to complete the programme. Further investigation of those in the 'too high' group found a trend in the data for the dropouts to be higher risk than programme completers (i.e. the highest risk of the most risky group). Furthermore, those within the 'too high' group who also had high criminogenic need scores seemed destined to be programme non-starters: none of the offenders in the 'too high' group with criminogenic need scores above 70 commenced the programme. This finding did not hold for those within the appropriately targeted group, however, indicating an interaction between risk and criminogenic need: an offender with high scores on both of these measures is likely to fail to commence whilst those with high scores relating to *either* risk *or* need can and do commence and complete the ETS programme. Indeed, those appropriately targeted offenders with high criminogenic need scores (over 50) who commenced the programme all completed it. Further research could usefully investigate the factors relating to programme commencement amongst this group.

Chapter seven reported that a third of all those who dropped out of their programme either could not or should not attend the ETS programme or were not able to attend because there was no programme place available to them. Indeed, the classification of programme dropouts into Stephens' (2003) typology of attrition revealed there to be subsets of the non-completer and non-starter groups. Similar to what had been found by Stephens and Turner (2004) and Briggs and Turner (2003), significant minorities of those sentenced to attend the ETS programme were later assessed as unsuitable for group or programme work or did not have a place on a programme either due to insufficient places or that the available places were inappropriate for their circumstances (e.g. employment, child care). The finding that unsuitable offenders were sentenced to ETS concurs with the research of Westmarland et al. (n.d.) and Turner (2006) and attests to the conclusions of Sussex Probation Area (2003) that more comprehensive training of pre-sentence report writers is required to reduce what might be termed 'administrative' attrition from programmes.

In addition to those who could not, should not, and were not able to attend due to a lack of places on the programme, a further third of programme dropouts failed to attend because they were in breach of their community orders prior to programme commencement. The remainder were otherwise compliant with their order but still failed to complete the programme. The reasons held within probation monitoring

systems for the non-attendance of this latter group were limited: some were removed from the programme due to behavioural issues, others were ill, or turned up late and hence were refused admission, but the majority were classified as non-attending with no reason provided as to why. More investigation utilising prospective qualitative research methods into the reasons why this latter group failed to attend or complete their programme would therefore be welcomed.

Investigation of the characteristics of the subsets of programme dropouts found that those who failed to commence or complete their programmes due to organisational reasons had profiles that were more similar to those of the programme completers than those of the programme non-completers and non-starters. When the dropout groups were purged of these individuals, therefore, those representing "true attrition" (Stephens, 2003, p.236) remained. This group were younger and higher risk, were more likely to have been subject to previous breach proceedings, commenced offending at a younger age, offended at a faster pace, and had more criminogenic needs than the other types of programme dropouts.

Given it is already known that programme dropouts are more criminogenic and at a greater risk of reconviction than programme completers, the finding that the subset of programme dropouts representing true attrition is characterised by even greater levels of criminality is enlightening. From a theoretical perspective, such findings raise the question that if risk, or indicators of risk, is able to differentiate not only programme completers, non-completers and non-starters but also subsets of programme dropouts relating to true attrition, has the risk principle (Andrews & Bonta, 2006) been violated? Beyko and Wong (2005) argue this to be the case. The quote below highlights their viewpoint in relation to sex offender treatment but can be applied equally to general offending behaviour programmes:

> A program that could provide services to offenders at all levels of risk should not have any reliable predictors of attrition that are related to offenders' sexual recidivism risk. On the other hand, a program that only caters to, for example, low-risk offenders would have predictors of attrition that are risk related (high risk offenders are more likely to drop out) if the program does not screen out all high-risk admissions (p. 377).

On first evaluation, therefore, it would appear that the risk principle has been violated: the risk of reconviction scores predicted programme attrition such that high risk offenders were more prone to dropout than appropriate or low risk offenders. However, further investigation shows it to be less straightforward than this. The finding that OGRS2 did not differ significantly between the appropriately targeted programme completers, non-completers and non-starters would attest that the targeting criteria do indeed respect the risk principle. It is hence the decisions made by correctional services officers to refer offenders whose risk of reconviction scores fall outside of these criteria that violates the risk principle.

However, the picture is even further complicated by the findings reported within chapter six. Three in ten of those offenders who are 'too high' risk for the ETS programme did manage to complete the programme. When this is combined with the knowledge that high risk offenders who complete programmes show the largest gains in relation to their treatment outcomes (Palmer et al., 2008, 2009) and in the absence of interventions specifically tailored to this group, it is perhaps understandable that correctional service officers are keen to pursue this possibility. As such, it would seem that more investigative work is needed to determine the factors associated with programme completion amongst high risk offenders. Chapter six was able to show that a combination of high risk and high need inevitably resulted in non-attendance at session one, but more work is needed to tease out the issues in this area.

In relation to the need principle (Andrews & Bonta, 2006), Beyko and Wong state:

> ...programs that only cater to treat offenders with certain criminogenic needs would have attrition predictors that are linked to the criminogenic needs that are not adequately addressed. For example, a program that is not equipped to deal with offenders with high levels of aggression would have attrition predictors linked to aggression, again, provided that the program does not have admission criteria that automatically screens out all aggressive offenders (p. 377).

Assessment of the need profiles of the programme dropouts presents a mixed picture. First, chapter three assessed the suitability (assessed as the level of cognitive deficit) and level of need (overall, criminogenic and non-criminogenic derived from the mini-OASys) of programme completers, non-completers and non-starters and concluded that there were no differences in these measures between the groups. However, differences were highlighted between the programme completers, non-completers and non-starters within chapter five in relation to the OASys total score. Whether the OASys total score should be used as a measure of need, however, is perhaps debateable. The OASys tool was developed to measure risk and need and whilst the separate need domains do contribute to the overall OASys score, they are combined with criminal history information to create what is generally known as a measurement of risk rather than need (Howard, 2006). As such, the separate criminogenic and non-criminogenic need scores derived from OASys data perhaps represent the needs of offenders better than the overall OASys score as the criminal history information is removed from their calculation. Examination of these scores, however, shows that the groups do differ in relation to criminogenic need with non-starters having more needs than the completers.

At first assessment, such a finding might lead to the conclusion that the ETS programme does not address the criminogenic needs of those sentenced to it and hence violates the need principle. However, when considered in conjunction with the above finding relating to the percentage of the sample who had OGRS2 scores above the recommended criteria and the finding in chapter three that OGRS2 scores and criminogenic need are positive correlated, it becomes likely that a similar conclusion can be reached in relation to the need principle as was reached in relation to the risk principle. Whilst the criteria set out by programme developers within programme manuals would seem to meet the principles of risk and need, the overriding of the risk principle by correctional staff introduces offenders to the programme whose needs are perhaps too great to be managed and addressed by a general offending behaviour programme designed for medium to high risk offenders. As such the inappropriate targeting of offenders serves to violate not only the risk principle but also the need principle of effective offender intervention.

The final principle of effective practice contained within the RNR theory of Andrews and Bonta (2006) is that of responsivity. In commenting on the relationship of this principle to the issue of attrition, Beyko and Wong (2005) state:

> A similar argument would apply to the responsivity dimension. A well designed and delivered treatment program should have very few "reliable" predictors for attrition since the risk, need, and responsivity characteristics of the participants are already adequately addressed by the program and these characteristics should not systematically predict attrition (p. 377).

Within this thesis and as discussed above, it has been found that (aside from the risk and need variables mentioned above) process factors also contribute to attrition. Whilst few of the factors studied in chapter six were found to differentiate programme completers from non-completers and non-starters, the findings relating to inappropriate targeting, insufficient places on programmes, a lack of flexibility of scheduling, and the sentencing of unsuitable offenders (due, for example, to mental health problems, an inability to cope in a group, or intellectual capacity issues) demonstrate a lack of responsivity. It would seem therefore that in order to be responsive to the needs of the offenders on probation a range of solutions are required. First and as mentioned above, training is needed throughout the correctional service to ensure that unsuitable referrals are not made to a general offending behaviour programme.

Second, there should be a consideration at the national level as to the appropriate provision for high risk offenders on a community sentence. The recommendations from the National Probation Directorate (2001) state that high risk offenders can be allocated to a general offending behaviour programme if it is sequenced with further interventions (as according to the risk principle these offenders require a greater level of intervention). Such interventions might include those specific to a type of offence or the generic cognitive skills booster programme, for example. However, if it is known that in sanctioning such treatment for high risk offenders that the majority will not commence the first of the interventions in the sequence, such practice could be deemed to be placing offenders in a situation where they are likely to fail and, hence, could be considered unethical. As failure to complete or commence the

intervention could result in the instigation of breach proceedings and the subsequent punishment arising from these, the advice given to probation services in relation to this issue should perhaps be reconsidered.

Third, if the above two issues were addressed it is possible that the current level of programme provision would be appropriate; fewer offenders would be sentenced to the general offending behaviour programmes and hence provision would not necessarily need to be increased. However, the level of accredited programme provision was not the sole reason why programme places were not available to offenders. Other reasons, such as the lack of flexibility of programmes to cater for employment patterns or child care responsibilities also need to be addressed by the probation service. In order to be truly responsive to such offender needs the probation service has to determine a way to ensure programmes are flexible enough to respond to such demands whilst maintaining the integrity of the treatment delivered. Such challenges are not easily overcome.

Future research

There are many avenues of further research stemming from the findings contained within this thesis. At relevant points within the above discussion some of these have already been highlighted. This section will highlight some ideas as to how this research could be taken forward and will suggest some projects that would usefully build on the findings presented here. Whilst this discussion is not exhaustive, it provides an indication as to possible lines of research enquiry. Some of these relate to more global questions, such as the continued investigation of the impact of attrition on reconviction outcomes, whilst others are more concerned with the minutiae of programme dropout: who is more likely to dropout, within which situation, and why.

Commencing with the more global issues, further work should be undertaken to tease out the impact of accredited programmes, and in particular the impact of attrition from these programmes, on reconviction outcomes. A number of differences between the programme completers, non-completers and non-starters have been highlighted within this thesis. Work should be undertaken to determine whether after statistically controlling for these variables, or after matching the experimental and comparison groups on these variables, programme completion/dropout has an effect on the subsequent reconviction of offenders. The ‘would do well anyway’ argument (Debidin &

Lovbakke, 2005) would proffer that the differences in reconviction rates that have been observed between the programme completer, non-completer, non-starter and comparison groups (Hatcher et al., 2008; Hollin et al., 2004, 2008; Palmer et al., 2007; McGuire et al., 2008) are a function of pre-existing differences between the groups on some unmeasured variable(s). Reconviction analyses which control (either statistically or by matching) for such differences therefore serve to test the influence of these variables; if differences in the outcome variable between the groups remain there is a greater chance that these differences are a function of programme completion or dropout rather than of self-selection.

Another body of research which would usefully feed into clinical practice relates to further investigation of the 'too high' group of offenders. Whilst the dropout rate amongst this group is high, previous research (Palmer et al., 2008) has indicated that the effect of programme completion amongst this group is large. Further research should therefore examine this group in more detail. In the absence of an alternative specific intervention for these offenders within the community services of England and Wales, it would be useful to understand the reasons for attrition (and completion) amongst this group and whether these can be attended to in a responsive manner in order to increase completion rates. Of course if this were possible, it would be necessary to monitor the outcomes of this group to ensure that the treatment gains continue to be as positive as those already indicated amongst this sample (Palmer et al., 2008). If these were not to persist once the completion rate increased, alternative interventions would need to be considered for high risk offenders within the community.

In relation to the treatment readiness theory and literature (Ward et al., 2004), it would be useful to combine the findings of this research with additional exploration of those factors associated with offender engagement in accredited programmes. As outlined in chapter one, Day and colleagues have stated that "a consequence of being 'ready for treatment' is the ability to not only attend rehabilitation programmes, but also to engage with the programme content, facilitators, and other group members" (Day et al., 2007, p. 22). In other words to be 'treatment ready' an offender first has to attend a programme and then has to engage with it and its associated processes. This thesis, in evaluating the non-starters against the non-completers and completers of programmes, has shed some light on the first of these conditions; it has, in part, been

concerned with those individual and organisational factors which impact on attendance (or not) at the programme to which an offender is sentenced. However, additional research assessing the correlates of engagement with programmes is warranted.

One way in which this could be undertaken would be to evaluate in more detail the differences between programme completers and non-completers. The act of non-completion may in itself be an indicator of low engagement with the programme. In the light of the findings reported above and in accord with the theory of treatment readiness, both individual and process factors should be examined. Additionally, research could usefully determine those programme completers who responded positively to the programme (i.e. those that engaged with it) and undertake a comparison of these with those who did not (i.e. those that did not engage). Within such research, positive indications of pre- to post-programme change or an absence of reconviction for a certain period of time could be taken as indicators of positive engagement with the intervention. Comparisons of those who responded positively could be made with those that did not in order to understand more about programme engagement and readiness for treatment. These groups could be studied quantitatively or qualitatively to determine the factors associated with treatment engagement and disengagement.

Finally, a prospective qualitative examination of programme completers, non-completers and non-starters would usefully add to the research base. To date, very few studies have undertaken to speak with offenders about their experiences of programme completion or dropout and those that have either have not involved programme completers and dropouts and/or have undertaken the research retrospectively. Of course, to interview programme completers, non-completers and non-starters after the event would provide useful data, however, the stories of the offenders are likely to be contaminated by the process with which they have been through. As such a prospective design which undertakes to interview the offenders pre- and post-programme would be a useful addition to the literature. Such a project could investigate the offenders' knowledge and expectations prior to the programme and then evaluate how these combine with personal, situational, and process factors to culminate in their post-programme status as a programme completer, non-completer or non-starter. To the

author's knowledge, there have been no such prospective qualitative investigations of programme attrition to date.

Critique of the research within this thesis

The author of this thesis has endeavoured to ensure that the data, data analysis, and information presented within this document are of the highest possible quality. However, it is the often the case when working within an applied field that problems are encountered in the course of the investigation which impact on the research and hence limit the conclusions that can be drawn.

Within this thesis, the research has relied exclusively on archival data or data collected for the purpose of monitoring offenders and their progress on their sentences. The advantage of obtaining and utilising data from existing databases is that it limits the demands that the research places on organisations, such as the probation service, and on the participants of the research. Additionally, where possible the aim was to use information that was routinely available to probation staff. In doing this any recommendations stemming from the research could be implemented into practice without the service having to collect any additional data on their clients.

As such, the data underpinning the research within this thesis was collected from the CRAMS, IAPS, OASys, and Offender Index databases. In using these databases, however, the research can only ever be as good as the recording of the data within them. Any limitations in data recording will hence translate into limitations of the research. For example, the Offender Index is a "database intended to contain all court disposals relating to standard list offences since 1963 in England and Wales" (Francis, Crosland, & Harman, 2002, p. 1). However, there are limitations to the use of this database as a research tool. Francis et al. (2002) outline the disadvantages as: "excludes Northern Ireland and Scotland, not a complete history for older offenders, delay in collection and processing of information, criminal histories may contain composite information on more than one individual, only standard list offences recorded, no cautions or warnings, no dates of offence – problem of pseudo-reconvictions" (p. 2). Such warnings highlight that research using such systems must, as a matter of course, accept that the data upon which the research is based may not be completely accurate. In the absence of alternative, easily accessible and/or more reliable methods of

assessing reconviction, however, and so as not to place too much of a research load on a service which is already stretched, the limitations of the various databases must be accepted. Where possible, and so as to limit inaccuracies within the data, thorough cleaning and cross-referencing of the information retrieved from probation databases was undertaken by the researcher.

A second issue with utilising existing data from probation databases is that the research is limited by the type of data that can be collected for analysis. Take, for example, the data used as an indicator of offender motivation to attend the ETS programme presented within chapter three. The data available to the researcher was collected by the probation service as part of the programme assessment process and consisted of the probation officers' ratings of their perceptions of how motivated to attend the programme the offenders were. This was rated on a Likert scale which ranged from one to five. Aside from the definitional issues relating to the concept of motivation (Drieschner et al., 2004), it is perhaps not ideal that an assessment of an internal concept such as motivation is made by someone who does not necessarily know the offender particularly well and when the relationship between the offender and the rater is perhaps such that it would be advantageous to the offender for the rater to believe that he or she is motivated to attend the programme. The quote previously used within the chapter one of this thesis resonates here: "Motivation to change is difficult to assess...because there are clear benefits to "appearing" willing to change" (Hanson & Bussiere, 1998, p. 349).

Similarly, the data used to determine the suitability of offenders for programmes utilised within chapter three has limitations. These data are collected by probation officers to inform their assessment of the cognitive deficits of the offender(s) under their supervision and hence their need for intervention. As such, and like the motivation scores, the data collection tool has not been developed for research purposes and hence has not been validated against other cognitive deficit assessments nor has it been assessed for inter-rater reliability. Hence, whilst instruction is provided, it is possible that different probation officers may score the same person differently. Further, the tool provides a crude indication of the level of cognitive deficit but does not discriminate those with different types of deficit; indeed two offenders may have identical scores but very different deficits. Such criticisms, however, could also be made of tools which are

utilised often in research: for example, the Hare's Psychopathy Checklist-Revised (2003) utilises a similar checklist approach alongside a pre-determined cut off which can result in two very different individuals both being labelled with the same term.

Problems such as those discussed above are inevitable, however, when research attempts to utilise tools which have been developed for other purposes. Another example within this research includes the use of the Home Office categorisation of offence types which results in an 'Other' category which inevitably includes a heterogeneous sample of different offenders. The aim of the research, however, was to determine correlates and factors which predict programme attrition using a pool of data available to probation areas. The reasoning behind this decision was to ensure that the outcomes of the research were practically useful to probation areas. It was also necessary due to the retrospective nature of the research; to collect the more reliable and valid information prospectively would necessitate a large resource input. However, it does have to be accepted that in making the decision to undertake a retrospective design utilising such tools, there is an inevitable loss of data quality which in turns impacts on the conclusions that can be drawn from the research.

Another limitation of this thesis relates to the sample sizes utilised within the analyses. Despite having reasonable samples of 193 participants within part one and 297 participants within part two, the sub-samples of programme completers, non-completers and non-starters are inevitably smaller. Further classification, such as that seen within chapters six and seven thus results in sub-samples containing few participants. As such, some of the analyses within this thesis were underpowered and, in a small number of cases, it was not possible to conduct the analyses due to insufficient participants. As such where the numbers have been low, the conclusions of the analyses should be regarded as tentative until further research either confirms or refutes them.

Another concern is whether the findings of this research are generalisable over location and time. Part two of this thesis was conducted entirely within one probation area and hence it is possible that the findings from this research may not be relevant to other probation areas. However, there are indications that portions of the research are generalisable. For example, the classification of the offenders into Stephens' typology of attrition groups in chapter seven revealed similar findings to those reported from West

Yorkshire Probation Area. In addition, the appropriate management and delivery of accredited programmes is specified within a centrally provided management manual and within centrally provided accredited programme training. Despite these safeguards, however, it is likely that the management and delivery of programmes is affected by differing offender profiles, different programme facilitators, and a multitude of other factors that vary between delivery sites. Likewise, this research represents a snapshot of the circumstances relevant to a certain period of time. Whilst the data within the two parts of the thesis present some overlapping findings from two different time points, it is possible that the rates and influences of attrition vary with time as guidance is altered, the programmes become more embedded within services, and practice is influenced by evidence (anecdotal or scientific). As such, replication of this research within alternative jurisdictions and under current conditions would be welcomed.

Conclusion

This thesis aimed to investigate the issue of attrition from community-based general offending behaviour programmes. In doing so, the research has evaluated the impact of this phenomenon on reconviction outcomes, has investigated the differences between programme completers, non-completers and non-starters, and has attempted classification of programme dropouts by the reasons for their failure to attend or complete the programme to which they were sentenced. The research has tentatively suggested the presence of a non-completion effect: that is, there is some evidence that failing to commence or complete the programme to which an offender is sentenced may have a detrimental effect on their reconviction outcomes. Additionally, the research has indicated that programme non-starters are distinct from programme completers: they commenced offending at a younger age, had more convictions as a juvenile, have committed offences at a faster pace and are more likely to have a history of breach. The research has also suggested that the non-completers are younger versions of the non-starters with a tendency to be more aggressive.

This thesis has shown that in line with treatment readiness theory (Ward et al., 2004) there is also an organisational element to attrition, however. It would seem that the practices undertaken in the targeting of offenders to accredited programmes may serve to violate the principles of risk and need. In addition, it would appear that

probation areas could do more to ensure that their delivery of accredited programmes is as responsive to the needs of offenders as possible. It is perhaps falls to the research community, with the cooperation of probation areas, to conduct further investigation within this field so as to accumulate the evidence needed to advise the correctional services on how their provision can be adapted to improve completion rates. The challenge for these services will then be to ensure that appropriate offenders are sentenced to programmes which are implemented in a manner responsive to the needs of these offenders and which permits appropriate flexibility to cater for the inevitable events that may, if permitted, prevent attendance on an accredited programme.

References

References:

Al-Attar, Z. (2003). Cognitive-behavioural programmes for persistent offenders: Evaluating effectiveness. *VISTA, 7,* 126-133.

Andrews, D. A. (1995). The psychology of criminal conduct and effective treatment in J. McGuire (Ed.) *What Works: Reducing Re-offending* (pp 35–62).Chichester: John Wiley & Sons,

Andrews, D. A. (2001). Principles of effective correctional programs. In L. L. Motiuk & R. C. Serin (Eds.) *Compendium 2000 on effective correctional programming* (pp 9-17). Ottawa: Correctional Service Canada.

Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct (4th ed.)*. Cincinnati, OH: Anderson Publishing.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006).The recent past and near future of risk and/or need assessment. *Crime and Delinquency, 52,* 7-27.

Andrews, D. A., & Dowden, C. (2006). Risk principle of case classification in correctional treatment: A meta-analytic investigation. *International Journal of Offender Therapy and Comparative Criminology*, *50*, 88-100.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P, & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28*, 369-404.

Antonowicz, D. H., & Ross, R. R. (1994). Essential components of successful rehabilitation programs for offenders. *International Journal of Offender Therapy and Comparative Criminology, 38*, 97-104.

Akers, R. L., Krohn, M. D., Lanza-Kaduce, L., & Radosevich, M. (1979). Social learning and deviant behaviour: A specific test of a general theory. *American Sociological Review, 44*, 636-655.

Babock, J. C., & Steiner, R. (1999). The relationship between treatment, incarceration, and recidivism of battering: A program evaluation of Seattle's coordinated community response to domestic violence. *Journal of Family Psychology, 13*, 46-59.

Baekeland, F., & Lundwall, L. (1975). Dropping out of treatment: A critical review *Psychological Bulletin, 82*, 738–783.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26.

Beyko, M. J. & Wong, S. C. P. (2005). Predictors of treatment attrition as indicators for program improvement not offender shortcomings: A study of sex offender treatment attrition. *Sexual Abuse: A Journal of Research and Treatment, 17*, 375-389.

Blanchette, K., & Brown, S. L. (2006). *The assessment and treatment of women offenders: An integrative perspective*. West Sussex, UK: Wiley.

Bloom, B., Owen, B. & Covington, S. (2003). *Gender responsive strategies: Research, practice and guiding principles for women offenders*. U.S. Department of Corrections, National Institute of Corrections.

Blud, L., & Travers, R. (2001). Interpersonal problem-solving skills training: A comparison of R&R and ETS. *Criminal Behaviour and Mental Heath, 11*, 251-261.

Blud, L., Travers, R., Nugent, F., & Thornton, D. (2003). Accreditation of offending behaviour programmes in HM Prison Service: "what works" in practice. *Legal and Criminological Psychology, 8,* 69-81.

Bodenhausen, G.V. (1990). Stereotypes and judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science*, *1* , 319–322.

Bonta, J., Wallace-Capretta, S., & Rooney, J., (2000). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior, 27*, 312-329.

Briggs, S., Gray, B., & Stephens, K. (2004). Offender literacy and attrition from the Enhanced Thinking Skills Programme. *VISTA, 8*, 73-80.

Briggs, S., & Turner, R. (2003). *Barriers to starting programmes: Second phase report*. Unpublished report. West Yorkshire Probation Area.

Brocato, J. & Wagner, E. F. (2008). Predictors of retention in an alternative-to-prison substance abuse treatment program. *Criminal Justice and Behavior, 35*, 99-119.

Browne, K. D., Foreman, L., & Middleton, D. (1998). Predicting treatment drop-out in sex offenders. *Child Abuse Review, 7*, 402-419.

Butzin, C. A., Saum, C. A., & Scarpitti, F. R. (2002). Factors associated with completion of a drug treatment court diversion programme. *Substance Use and Misuse, 37,* 1615-1633.

Cann, J., Falshaw, L., Nugent, F., & Friendship, C. (2003). *Understanding what works: Accredited cognitive skills programmes for adult men and young offenders*. Home Office Research Findings No. 226. London: Home Office.

Canton, R., & Hancock, D. (2007). Dictionary of probation and offender management. Cullomption, Devon: Willan Publishing.

Casey, S., Day, D., Howells, K., & Ward, T. (2007). Assessing suitability for offender rehabilitation: Development and validation of the Treatment Readiness Questionnaire. *Criminal Justice and Behavior, 34*, 1427-1440.

Chopourian, G. (2003). *Accredited programmes attrition: Focus upon Think First in West Midlands Probation Area.* Unpublished report.

Chitty, C. (2005). The impact of corrections on re-offending: Conclusions and the way forward. In G. Harper and C. Chitty, *The impact of corrections on re-offending: a review of 'what works'. Home Office Research Study 291*, 3rd edition (pp. 75-82). London: Home Office.

Chu, S. (2003). *ThinkFirst programme attrition: The impact of case management.* Unpublished report. South Yorkshire Probation Area.

Clark, D. (2000). *Theory manual for Enhanced Thinking Skills*. Unpublished report. Prepared for the Joint Prison Probation Service Accreditation Panel.

Clarke, A., Simmonds, R., & Wydall, S. (2004). Delivering cognitive skills programmes in prison: A qualitative study. Home Office Research Findings No. 242. London: Home Office.

Colegrave, N., & Ruxton, G. D. (2003). Confidence intervals are a more useful complement to nonsignificant tests than are power calculations. *Behavioral Ecology, 14*, 446-447.

Colledge, M., P. Collier and S. Brand (1999) *Programmes for offenders: Guidance for evaluators*. CRP Guidance Note 2. London: Home Office.

Copas, J. & Marshall, P. (1998). The offender group reconviction scale: a statistical reconviction score for use by probation officers. *Applied Statistics, 47*, 159-171.

Craig, A. R., Franklin, J. A. & Andrews, G. (1984). A scale to measure locus of control of behaviour. *British Medical Journal of Medical Psychology,* 57, 173-180.

Craissati, J., & Beech, A. (2001). Attrition in a community treatment programme for child sexual abusers. *Journal of Interpersonal Violence, 16*, 205-221.

Criminal Justice Act (2003). Retrieved March 20, 2009 from http://www.opsi.gov.uk/acts/acts2003/ukpga_20030044_en_1

Dancey, C., & Reidy, J. (2004). *Statistics without maths for psychology*. London, UK: Prentice Hall

Davies, K., Lewis, J., Byatt, J., Purvis, E., & Cole, B. (2004). *An evaluation of the literacy demands of general offending behaviour programmes*. Home Office Research Findings 233. London: Home Office.

Dawson, P., Walmsley R. K., & Debidin, M. (2005). *A process evaluation of the cognitive skills booster programme initial roll-out in 14 probation areas and 12 prisons.* Home Office Online Report 41/05. London: Home Office.

Day, A., Howells, K., Casey, S., Ward, T., & Birgden, A. (2007). Treatment readiness: An overview of Australasian work. *Issues in Forensic Psychology, 7,* 21-25.

De Leon, G., & Jainchill, N. (1986). Circumstance, motivation, readiness and suitability as correlates of treatment tenure. *Journal of Psychoactive Drugs, 18,* 203–208.

De Leon, G., Melnick, G., Thomas, G., Kressel, D., & Wexler, H. K. (2000). Motivation for treatment in a prison-based therapeutic community. *American Journal of Drug and Alcohol Abuse, 26,* 33–46.

Debidin, M. & Lovvbakke, J. (2005). Offending behaviour programmes in prison and probation. In G. Harper and C. Chitty, *The impact of corrections on re-offending: a review of 'what works'. Home Office Research Study 291*, 3rd edition (pp. 31-51). London: Home Office.

Digman, J. M. (1990). Personality Structure: emergence of the five factor model. *Annual Review of Psychology*, *41*, 417-440.

Dodd, N. J. (1998). Applying psychology to the reduction of insurance claim fraud. *Insurance Trends, 18*, 11-16

Dowden, C. (1998). *A meta-analytic examination of the risk, need and responsivity principles and their importance within the rehabilitation debate.* Unpublished thesis. Carleton University: Ottawa, Ontario, Canada.

Dowden, C. & Andrews, D. A. (1999a). What works for female offenders: A meta analytic review. *Crime and Delinquency, 45*, 438-452.

Dowden, C. & Andrews, D. A. (1999b). What works in young offender treatment: A meta analysis. *Forum on Corrections Research, 11*, 21-24.

Dowden, C. & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology and Criminal Justice, 42*, 327-342.

Drieschner, K. H., Lammers, S. M. M., & van der Staak, C. P. R. (2004). Treatment motivation: An attempt for clarification of an ambiguous concept. *Clinical Psychology Review, 23*, 1115-1137.

Eysenck, S. B., and Eysenck, H.J., (1978). Impulsiveness and venturesomeness, their position in a dimensional system of personality description. *Psychological Reports, 43,* 1247-1255.

Farrington, D.P. (1992). Explaining the beginning, progress, and ending of antisocial behaviour from birth to adulthood. In Joan McCord (ed.), *Facts, frameworks and forecasts: Advances in criminological theory, Vol. 3.* New Brunswick, N.J.: Transaction.

Farrington, D. P. (2003). Developmental and life- course criminology: Key theoretical and empirical issues – The 2002 Sutherland Award Address. *Criminology, 41,* 221-255.

Farrington, D.P., Gottfredson, D. C., Sherman, L. W., & Welsh, B.C. (2002). The Maryland Scientific Methods Scale. In L.W. Sherman, D.P. Farrington, B.C. Welsh & D.L. MacKenzie (Eds.) *Evidence-based crime prevention*. pp. 3–21. London: Routledge.

Farrington, D.P. & West, D.J. (1993). Criminal, penal and life histories of chronic offenders: Risk and protective factors and early identification. *Criminal Behaviour and Mental Health*, 3, 492-523.

Field, A. (2005). *Discovering statistics using SPSS. Second edition.* London: Sage.

Fiore, M.C., Bailey, W.C., Cohen, S.J., Dorfman, S.F., Goldstein, M.G., et al. (1996). *Smoking cessation. Clinical practice guideline No 18. Publication No 96–0692.* Agency for Health Care Policy and Research, US Department of Health and Human Services: Rockville, MD.

Francis, B., Crosland, P., & Harman, J. (2002). The Police National Computer and the Offenders Index: can they be combined for research purposes? Home Office Research Findings 170. London: Home Office.

Francis, B., Harman, J., & Humphreys, L. (2005). Predicting reconviction rates in Northern Ireland. Northern Ireland Office Research and Statistics Bulletin 7/2005. Belfast: Northern Ireland Office.

Freedman, B. J., Rosenthal L., Donahoe, C. P., Schlundt, D. G., & McFall, R. M. (1978). A social behavioural analysis of skill deficits in delinquent and non-delinquent adolescent boys. *Journal of Consulting and Clinical Psychology, 44*, 1448-62.

Friendship, C., Blud, L., Erikson, M., Travers, R., & Thornton, D. (2003). Cognitive-behavioural treatment for imprisoned offenders: An evaluation of HM Prison Service's cognitive skills programmes. *Legal and Criminological Psychology*, *8*, 103-114.

Friendship, C., Thornton, D., Erikson, M., & Beech, A. (2001). Reconviction: A critique and comparison of two main data sources in England and Wales. *Legal and Criminological Psychology, 6,* 121-130.

Frude, N., Honess, T. M. and Maguire, M. (1994). *Crime-PICS II: A psychometric tool for measuring attitude change in probation clients*. Swansea, UK: Michael and Associates.]

Geer, T. M., Becker, J. V., Gray, S. R., Krauss, D. (2001). Predictors of treatment completion in a correctional sex offender treatment program. *International Journal of Offender Therapy and Comparative Criminology, 45,* 302-313.

Gill, E. J. (2004). *Attrition or completion on the Think First programme: Offender perspectives*. Unpublished thesis (MSc Forensic Psychology). Manchester Metropolitan University.

Girard, L. (1999). *The Level of Service Inventory–Ontario Revision: Risk/need assessment and recidivism*. Unpublished doctoral dissertation, University of Ottawa, Ontario, Canada.

Gondolf, E. W. (2001). Limitations of experimental evaluation of batterer programs. *Trauma, Violence and Abuse, 2,* 79-88.

Gossop, M., Marsden, J., Stewart, D., & Rolfe, A. (2000). Reductions in acquisitive crime and drug use after treatment of addiction problems: 1-year follow-up outcomes. *Drug and Alcohol Dependence, 58,* 165-172.

Gough, H. (1960). Theory and measurement of socialisation. *Journal of Consulting Psychology 24*, 23-30.

Hanson, R. K. & Bussiere, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. Journal of Consulting and Clinical Psychology, 66, 348-362.

Hare, R. D. (2003). *The Psychopathy Checklist—Revised, 2nd ed.* Toronto: Multi-Health Systems.

Harper, G. & Chitty, C. (Eds.) (2005). *The impact of corrections on re-offending: A review of "what works".* Home Office Research Study 291 (2nd ed.). London: Home Office.

Hatcher, R. M., Palmer, E. J., McGuire, J., Hounsome, J. C., Bilby, C. A. L. & Hollin, C. R. (2008). Aggression Replacement Training with adult male offenders within community settings: A reconviction analysis. *Forensic Psychiatry and Psychology, 19,* 517-532.

Hedderman, C. (2007). Past, present and future sentences: What do we know about their effectiveness? In L. Gelsthorpe and R. Morgan (Eds.) *Handbook of Probation*. Chichester: John Wiley & Sons.

Herrick, E. (1991 December). The prison literacy connection. *Corrections Compendium Vol. XVI.*

Hazeltine, M., Walker, C., Nickells, K., & Gregg, S. (2003). *Study of attrition from the Think First programme in the Bedfordshire Probation Area.* Unpublished report. Bedfordshire Probation Area.

Heilburn, K., Dematteo, D., Fretz, R., Erickson, J., Yashuhara, K., & Anumba, N. (2008). How "specific" are gender-specific rehabilitation needs? An empirical analysis. *Criminal Justice and Behavior, 35*, 1382-1397.

Heinsman, D.T., & Shadish, W.R. (1996). Assignment methods in experimentation: When do nonrandomized experiments approximate answers from randomized experiments? *Psychological Methods, 1,* 154-169.

Hollin, C. R. (1995). The meaning and implications of 'programme integrity'. In J. McGuire (Ed.), *What works: Reducing reoffending: Guidelines from research and practice.* Chichester, UK: John Wiley & Sons.

Hollin, C. R. (2006). Evidence-based practice, manuals and evaluation. In C. R. Hollin and E. P. Palmer (Eds.), *Offending behaviour programmes: Development, application and controversies* (pp 33-67). Chichester, UK: John Wiley & Sons.

Hollin C. R. (2008). Evaluating offending behaviour programmes: Does only randomization glister? *Criminology and Criminal Justice, 8*, 89-106.

Hollin, C. R., McGuire, J., Hounsome, J., Hatcher, R., Bilby, C., & Palmer, E. J. (2008). Cognitive skills offending behaviour programmes in the community: A reconviction analysis. *Criminal Justice and Behavior, 35,* 269-283.

Hollin, C. R., McGuire, J., Palmer, E., Bilby, C., Hatcher, R**.,** & Holmes, A. (2002a). *Introducing Pathfinder programmes in the Probation Service*. Home Office Research Findings 177. London : Home Office.

Hollin, C. R., McGuire, J., Palmer, E., Bilby, C., Hatcher, R., & Holmes, A. (2002b). *Introducing Pathfinder programmes in the Probation Service: An interim report. Home Office Research Study 247. London: Home Office.*

Hollin, C. R., & Palmer, E. J. (2006). *Offending behaviour programmes: Development, application and controversies.* Chichester: John Wiley & Sons.

Hollin, C. R. & Palmer, E. J. (2009). *Cognitive skills programmes for offenders. Psychology, Crime and Law, 15, 147-164.*

Hollin, C. R., Palmer, E., McGuire, J., Hounsome, J., Hatcher, R**.,** Bilby, C., & Clark, C. (2004). *Pathfinder programmes in the Probation Service: A retrospective analysis*. Home Office Online Report 66/04. London : Home Office.

Hollis, V. (2007). *Reconviction analysis of programme data using Interim Accredited Programmes Software (IAPS)*. London, UK: National Offender Management Service.

Home Affairs Committee. (2007, June 15). *Young black people and the criminal justice system*. *Volume I.* London: Home Affairs Committee. Retrieved March 22, 2009, from http://www.publications.parliament.uk/pa/cm200607/cmselect/cmhaff/181/181i.pdf

Homel, P., Nutley, S., Webb, B., & Tilley, N. (2004). *Investing to deliver: Reviewing the implementation of the UK Crime Reduction Programme.* Home Office Research Study 281. London: Home Office.

Home Office (n.d.). *Think First: Programme outline*. London: Home Office.

Home Office (2000). *What Works: Implementation of Accredited programmes 2. Probation Circular 96/2000.* London: Home Office.

Home Office (2003a, September 16). *Working with offenders: Offending behaviour programmes*. Retrieved September 29, 2008, from http://www.crimereduction.homeoffice.gov.uk/workingoffenders/workingoffenders3.htm

Home Office (2003b, September 26). *Working with offenders: What works: Accreditation a summary.* Retrieved March 31, 2009, from

http://www.crimereduction.homeoffice.gov.uk/workingoffenders/workingoffenders13.htm

Howard, P. (2006). *The Offender Assessment System: an evaluation of the second pilot.* Home Office Research Findings 278. London: Home Office.

Howard, P., Francis, B., Soothill, K., & Humphreys, L. (2009). *OGRS 3: The revised Offender Group Reconviction Scale. Research summary 7/09.* London: Ministry of Justice.

Howells, K., & Day, A. (2006). Affective determinants of treatment engagement in violent offenders. *International Journal of Offender Therapy and Comparative Criminology, 50*, 174-186.

Howells, K. & Day, A. (2007). Readiness for treatment in high risk offenders with personality disorders. *Psychology, Crime and Law, 13*, 47-56.

Izzo, R. L., & Ross, R. R. (1990). Meta-analysis of rehabilitation programmes for juvenile delinquents. *Criminal Justice and Behavior, 17,* 134-142.

Kemshall, H., & Canton, R. (2002). *The effective management of programme attrition: A report for the National Probation Service (Welsh Region).* Unpublished report: National Probation Service.

Kolton, D. J. C. (2004). Active encouragement: Perceived coercion into correctional treatment. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64 (8-B),* 4045.

Kotrlik, J. W., & Williams, H. A. (2003). The incorporation of effect size in information technology , learning and performance research. *Information Technology, Learning and Performance Journal, 21*, 1-7.

Langevin, R. (2006). Acceptance and completion of treatment among sex offenders. *International Journal of Offender Therapy and Comparative Criminology, 50*, 402-417.

Levine, M., & Ensom, M. H. H. (2001). Post hoc power analysis: An idea whose time has passed? *Pharmacotherapy, 21,* 405-409.

Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. Cook, D. Cooper, H. Corday, H. Hartman, L Hedges, R. Light, T. Louis, & F. Mosteller (Eds.) Meta-analysis for explanation: A casebook (pp. 83-127). New York, NY: Russell Sage Foundation.

Lipsey, M. W. (1999). Can rehabilitative programs reduce the recidivism of juvenile offenders? An inquiry into the effectiveness of practical programs. *Virginia Journal of Social Policy and the Law, 6*, 611-641.

Lipton, D. S., Thornton, D., McGuire, J., Porporino, F., & Hollin, C. R. (2000). Program accreditation and correctional treatment. *Substance Use and Misuse, 35*, 1705-1734.

Lowenkamp, C. T., Latessa, E. J., & Holsinger, A. M., (2006). The risk principle in action: What we have learned from 13,676 offenders and 97 correctional programs. *Crime and Delinquency, 52*, 77-93.

Lowenkamp, C., Latessa, E., & Lemke, R. (2006, February 21). Evaluation of Ohio's RECLAIM funded programs, community corrections facilities and DYS facilities: FY 2002. CCF Supplementary Report. Retrieved September 24, 2008, from http://www.uc.edu/criminaljustice/ProjectReports/CCF_Evaluation_Final_2006.pdf

Lowenkamp, C., Latessa, E., & Smith, P. (2006). Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology and Public Policy, 5*, 575-594.

May, C.P., Hasher, L., & Stoltzfus, E. R. (1993). Optimal time of day and the magnitude of age differences in memory. *Psychological Science, 4*, 326-330.

McConaghy, N. (1999). Methodological issues concerning evaluation of treatment for sexual offenders: Randomization, treatment dropouts, untreated controls, and

within-treatment studies. *Sexual Abuse: A Journal of Research and Treatment, 11*, 183-193.

McGuire, J. (2005). The Think First programme. In M. McMurran & J. McGuire (Eds), *Social problem solving and offending: Evidence, evaluation and evolution* (pp. 183-206). Chichester: John Wiley & Sons.

McGuire, J. (2006). General offending behaviour programmes: Concept, theory and practice. In C. R. Hollin and E. J. Palmer (Eds.) *Offending behaviour programmes: Development, application, and controversies*. Chichester: John Wiley & Sons.

McGuire, J., Bilby, C., Hatcher, R. M., Hollin, C. R., Hounsome, J., Palmer, E. J. (2008). Evaluation of structured cognitive-behavioural treatment programmes in reducing criminal recidivism. *Journal of Experimental Criminology, 4,* 21-40.

McGuire, J., & Hatcher, R. (2002). Offence-focused problem-solving: Preliminary evaluation of a cognitive skills program. *Criminal Justice and Behavior, 28*, 564-587.

McMahon, G., Hall, A., Hayward, G., Hudson, C., & Roberts, C. (2004). Basic skills programmes in the probation service: An evaluation of the basic skills pathfinder. Home Office Research Findings No. 203. London: Home Office.

McMurran, M., & McCulloch, A. (2007). Why don't offenders complete treatment? Prisoners' reasons for non-completion of a cognitive skills programme. *Psychology, Crime and Law, 13*, 345-354.

McMurran, M., & Theodosi, E. (2007). Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime and Law, 13*, 333-344.

Merrington, S. & Stanley, S., (2007). Effectiveness: Who counts what? In L. Gelsthorpe and R. Morgan (Eds.) *Handbook of Probation*. Chichester: John Wiley & Sons.

Miner, M., & Dwyer, S. (1995). Analysis of dropouts from outpatient sex offender treatment. *Journal of Psychiatry & Human Sexuality*, *7,* 77–93.

Ministry of Justice (2007). *National standards for the management of offenders: Standards and implementation guidance 2007*. London, UK: Ministry of Justice.

Moffitt, T. E. (1993). Life-course-persistent and adolescence-limited antisocial behavior: A developmental taxonomy. *Psychological Review, 100,* 674-701.

Moore, D., Bergman, B., & Knox, P. (1999). Predictors of sex offender treatment completion. *Journal of Child Sexual Abuse*, *7,* 73–88.

Mosher, C., & Phillips, D. (2001, October 3). Program evaluation of the Pine Lodge pre-release residential therapeutic community for women offenders in Washington State. Washington, DC: National Institute of Justice. Retrieved September 23, 2008, from http://www.ncjrs.gov/pdffiles1/nij/grants/196670.pdf

National Audit Office (2001). *The implementation of the National Probation Service information systems strategy.* London: Her Majesty's Stationary Office.

National Offender Management Service (2007). *Accredited offending behaviour programmes: Performance improvement standards manuals.* London: National Offender Management Service.

National Management Manual (2000). *Effectively delivering accredited programmes in the community.* London: Home Office.

National Probation Service (2003a). About us: Our aims. Retrieved April 2, 2009, from http://www.probation.homeoffice.gov.uk/output/page18.asp

National Probation Service (2003b). *OASys: The new offender assessment system: Important information for sentencers*. London, UK: National Probation Service.

National Probation Directorate (2001). *Interim accredited programme software (IAPS): Data collection and reporting requirements.* Probation Circular PC 32:2001. London: Home Office.

National Probation Directorate (2004). *General offending behaviour/cognitive skills programmes: Evaluation manual and scoring supplement*. London, UK: National Probation Directorate.

National Statistics (2006, December 15). Focus on ethnicity and identity. Retrieved March 22, 2009, from http://www.statistics.gov.uk/focuson/ethnicity/

National Statistics (2008, September 26). *Crime: gender*. Retrieved September 29, 2008, from http://www.statistics.gov.uk/CCI/nugget.asp?ID=1661

Noble, J. A. (2006). Cultural exposure and attitudes toward racial diversity in Caucasians, Monoracials and Multiracials. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66(7-B),* 4000.

Nooney, K. (2004). *ThinkFirst offending behaviour programme: Analysis of programme short term effects from standardised self-reports.* Report for National Probation Directorate.

Nunes, K. L., & Cortoni, F. (2006a). *The heterogeneity of treatment non-completers.* Toronto, Canada: Correctional Services Canada.

Nunes, K. L. & Cortoni, F. (2006b). *Estimating risk of dropout and expulsion from correctional programs.* Toronto, Canada: Correctional Services Canada.

Ong, G., Al-Attar, Z., Roberts, C. and Harsent, L. (2003) *'Think First': An accredited community-based cognitive-behavioural programme in England and Wales. Findings from the prospective evaluation in three Probation Areas*. Unpublished report. University of Oxford.

Palmer, E. J., McGuire, J., Hatcher, R. M., Hounsome, J., Bilby, C. A. L., & Hollin, C. R. (2009). Allocation to offending behavior programmes in the English and Welsh Probation Service. *Criminal Justice and Behavior, 36*, 909-922.

Palmer, E. J., McGuire, J., Hatcher, R. M., Hounsome, J., Bilby, C. A. L., & Hollin, C. R. (2008). The importance of appropriate allocation to offending behaviour

programmes. *International Journal of Offender Therapy and Comparative Criminology, 52,* 206-221.

Palmer, E. J., McGuire, J., Hounsome, J., Hatcher, R., Bilby, C., & Hollin, C. R. (2007). Offending behaviour programmes in the community: The effects on reconviction of three programmes with adult male offenders. *Legal and Criminological Psychology, 12,* 251-264.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage.

Pelissier, B., Camp, S.D., & Motivans, M. (2003). Staying in treatment: How much difference is there from prison to prison? *Psychology of Additive Behaviours, 17*, 134-141.

Perry, D., & Johnson, C. (2000). *What works strategy for the probation service*. London: What Works Strategy Group.

Petros, T.V., Beckwith, B.E., & Anderson, M. (1990). Individual differences in the effects of time of day and passage difficulty on prose memory in adults. *British Journal ofPsychology*, *81*, 63–72.

Ponterotto, J. G., Burkard, A., Rieger, B. P., Grieger, I, D'Onofrio, A., Dubuisson, A., Heenehan, M., Millstein, B., Parisi, M., Rath, J. F., & Sax, G. (1995). Development and initial validation of the Quick Discrimination Scale (QDI). *Educational and Psychological Measurement, 55*, 1016-31.

Ponterotto, J. G., Potere, J. C., & Johansen, S. A. (2002). The Quick Discrimination Index: Normative data and user guidelines for counseling researchers. *Journal of Multicultural Counseling and Development, 30*, 192-207.

Quay, H. C. (1987). Institutional treatment. In H. C. Quay (Ed.). *Handbook of Juvenile Delinquency.* New York: Wiley.

Rankin, C. E. (2005). Illiterate prisoners? Myths and empirical realities. *Journal of Offender Rehabiliation, 41*, 43-55.

Raynor, P. (2003). Evidence-based probation and its critics. *Probation Journal, 50*, 334-345.

Raynor, P. (2004). The Probation Service 'Pathfinders': Finding the path and losing the way? *Criminal Justice, 4*, 309-325.

Raynor, P. (2008). Community penalties and Home Office research: On the way back to nothing works? *Criminology and Criminal Justice,* 8, 73-89.

Raynor, P. & Vanstone, M. (2007). Towards a correctional service. In L. Gelsthorpe and R. Morgan (Eds.) *Handbook of Probation*. Chichester: John Wiley & Sons

Rea, L. M., & Parker, R. A. (1992). *Designing and conducting survey research*. San Francisco: Jossey-Bass.

Redondo, S., Sanchez-Meca, J., & Garrido, V. (1999). The influence of treatment programmes on the recidivism of juvenile and adult offenders: An European meta-analytic review. *Psychology, Crime and Law, 5,* 251-278.

Roberts, C. (2004). An early evaluation of a cognitive offending behaviour programme ('Think First') in probation areas. *VISTA, 8*, 130-136.

Robinson, D. (1995). The impact of cognitive skills training on post-release recidivism among Canadian federal offenders. Ottawa: Correctional Service of Canada.

Robinson, D., Grossman, M., & Porporino, F. (1991). *Effectiveness of the Cognitive Skills Training Programme: From pilot to national implementation.* Report B-07. Ottawa: Correctional Service of Canada.

Ross, R. R., & Fabiano, E. A., (1985). *Reasoning and Rehabilitation: Manual*. Ottawa: AIR Training & Associates.

Ross, R. R., Fabiano, E. A., & Ewles, C. D., (1988) Reasoning and Rehabilitation. *International Journal of Offender Therapy and Comparative Criminology, 32*, 29-36

Ryan, T.A. (1990) *Literacy training and the reintegration of offenders*. National Criminal Justice Reference Service. NCJ number 127211.

Ryan, R. M., Plant, R. W., & O'Malley, S. S. (1996). Initial motivations for alcohol treatment: Relations with patient characteristics, treatment involvement, and dropout. *Addictive Behaviors, 20,* 279–297.

Ryan, L., Hatfield, C., & Hofstetter, M. (2002). Caffeine reduces time-of-day effects on memory performance in older adults. *Psychological Science, 13,* 68-71.

Seager, J. A., Jellicoe, D., & Dhaliwal, G. K. (2004). Refusers, dropouts, and completers: Measuring sex offender treatment efficacy. *International Journal of Offender Therapy and Comparative Criminology, 46*, 600-612.

Serin, R. (1998). Treatment responsivity, intervention and reintegration: A conceptual model. *Forum on Corrections Research, 10*, 29-32.

Serin, R. C., Gobeil, E., & Preston, D., L. (2009). Evaluation of the Persistently Violent Treatment Program. *International Journal of Offender Therapy and Comparative Criminology, 53*, 57-73.

Serin, R., & Kennedy, S. (1997). *Treatment readiness and responsivity: Contributing to effective correctional programming.* Research report, Correctional Services Canada.

Scott, K. L. (2004). Stage of change as a predictor of attrition among men in a batterer treatment program. *Journal of Family Violence, 19,* 37-47.

Shaw, T., Herkov,M., & Greer, R. (1995). Examination of treatment completion and predicted outcome among incarcerated sex offenders. *Bulletin of the Academy of Psychiatry and Law*, *23,* 35–41.

Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J. E., Reuter,P., & Bushway, S. D. (1997). *Preventing crime: What works, what doesn't and what's promising.* Washington, DC: Department of Justice, National Institute of Justice.

Silagy, C., Mant, D., Fowler, G., & Lancaster, T. (1999). Nicotine replacement therapy for smoking cessation. *The Cochrane Library, 1*. Update Software, Oxford.

Simpson, D. D., & Joe, G. W. (1993). Motivation as a predictor of early dropout from drug abuse treatment. *Psychotherapy, 30,* 357–368.

Simpson, D. D., Joe, G. W., & Rowan-Szal, G. A. (1997). Drug abuse treatment retention and process effects on follow-up outcomes. *Drug & Alcohol Dependence, 47,* 227–235.

Social Exclusion Unit. (2002). *Reducing offending by ex-prisoners*. Report by the Social Exclusion Unit. London: Cabinet Office.

Spivack, G., & Levine, M. (1963). *Self-regulation in acting-out and normal adolescents*. Report M-453l. Washington, DC: National Institute of Mental Health.

Stark, M. J. (1992). Dropping out of substance abuse treatment: A clinically oriented review. *Clinical Psychology Review, 12*, 93-116.

Stephens, K. (2003). What is programme attrition? Towards an action-oriented typology. *VISTA, 7,* 233-240.

Stephens, K., & Turner, R. (2004). Barriers to programme start: Findings from an examination of case notes on a sample of required to attend programmes. *VISTA, 8*, 18-25.

Stewart-Ong, G., Harsent, L., Roberts, C., Burnetts, R., & Al-Attar, Z. (2004). *What works: NPD: Programme evaluation findings.* London: National Probation Service for England and Wales.

Stewart, D. Gossop, M. Marsden, J. & Rolfe, A. (2000). Drug misuse and acquisitive crime among clients recruited to the National Treatment Outcome Research Study (NTORS). *Criminal Behaviour and Mental Health, 10*, 10-20.

Sussex Probation Area (2003). Accredited programmes: Reducing attrition, increasing engagement. Unpublished report, Sussex Probation Area.

Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistics,* 5th ed. Boston: Allynn and Bacon.

Taylor, R. (1999). *Predicting reconvictions for sexual and violent offences using the revised offender group reconviction scale.* Research Findings No. 104. London, UK: Home Office Research Development and Statistics Directorate.

Thomas, L. (1997). Retrospective power analysis. *Conservation Biology, 11*, 276-280.

Turner, R. (2006). *Developing understanding of accredited programmes completions: The role of case-management and barriers to completion.* Unpublished report, West Yorkshire Probation Area.

Underdown, A. (2001). Making 'What Works' work: challenges in the delivery of community penalties. In A. Bottoms, L. Gelsthorpe, S. Rex (Eds.) *Community penalties: change and challenges* (pp 117-125). Chichester: John Wiley & Sons.

Van Der Zanden, B. P., Dijkgraaf, M. G. W., Blanken, P. Van Ree, J. M., & Van Den Brink, W. (2007). Patterns of acquisitive crime during methadone maintenance treatment among patients eligible for heroin assisted treatment. *Drug and Alcohol Dependence, 86,* 84-90.

van Voorhis, P., Spruance, L. M., Ritchey, P. N., Listwan, S. J., & Seabrook, R. (2004). The Georgia cognitive skills experiment: A replication of Reasoning and Rehabilitation. *Criminal Justice and Behavior, 31*, 282–305.

Walters, G. D. (2002). Current and historical content scales for the psychological inventory of criminal thinking styles. *Legal and Criminological Psychology, 7*, 73-86.

Ward, T., Day, A., Howells, K., & Birgden, A., (2004). The multifactor offender readiness model. *Aggression and Violent Behavior, 9*, 645-673.

Westmarland, N., Hester, M., Reid, P., Coulson, S., & Hughes, J., (n.d.). *An investigation into the factors associated with attrition in the Northumbria probation Think First programme.* Unpublished report: International Centre for the Study of Violence and Abuse, University of Sunderland.

Whitehead, J. T. & Lab, S. P. (1989). A meta-analysis of juvenile correctional treatment. *Journal of Research in Crime and Delinquency, 26*, 276-295.

Wilson, S. Attrill, G, & Nugent, F. (2003). Effective interventions for acquisitive offenders: An investigation of cognitive skills programmes. *Legal and Criminological Psychology, 8*, 83-101.

Wolfgang, M. E., Figlio, R. M., & Sellin, T. (1972). *Delinquency in a birth cohort.* Chicagom Ill.: University of Chicago Press.

Wood, J. (2005). Non attendance on the Think First programme: Offenders' perspective. *VISTA, 9*, 104-111.

Wormith, J. S., & Olver, M. E. (2002). Offender treatment attrition and its relationship with risk, responsivity, and recidivism. *Criminal Justice and Behavior, 29*, 447-471.

Young, D. (2002). Impacts of perceived legal pressure on retention in drug treatment. *Criminal Justice and Behavior, 29*, 27-55.

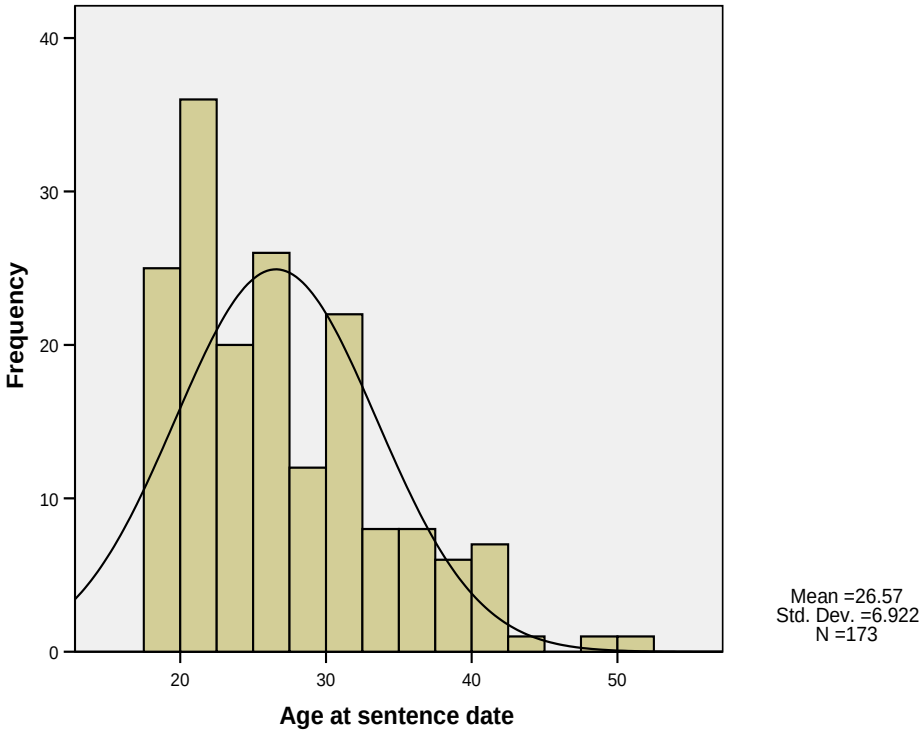
Yuan, K., & Maxwell, S. (2005). On the post hoc power in testing mean differences. *Journal of Educational and Behavioral Statistics, 30,* 141-167.

Zanis, D. A., Mulvaney, F., Coviello, D., Alterman, A. I., Savitz, B., & Thomson, W. (2003). The effectiveness of early parole to substance abuse treatment facilities on 24-month criminal recidivism. *Journal of Drug Issues , 33*, 223-236.

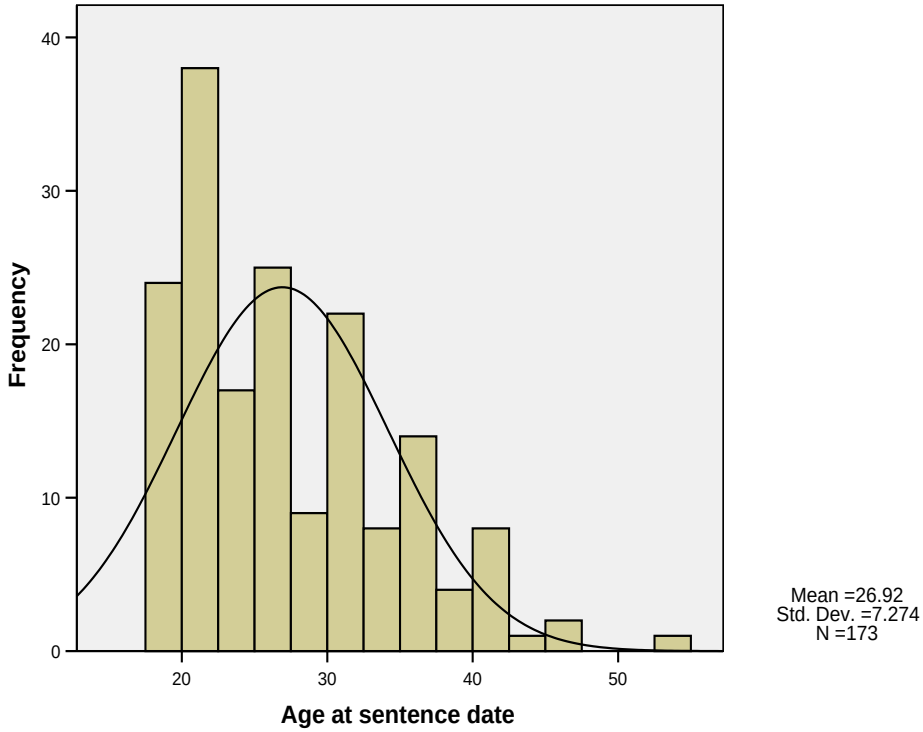
Appendices

Appendix A:

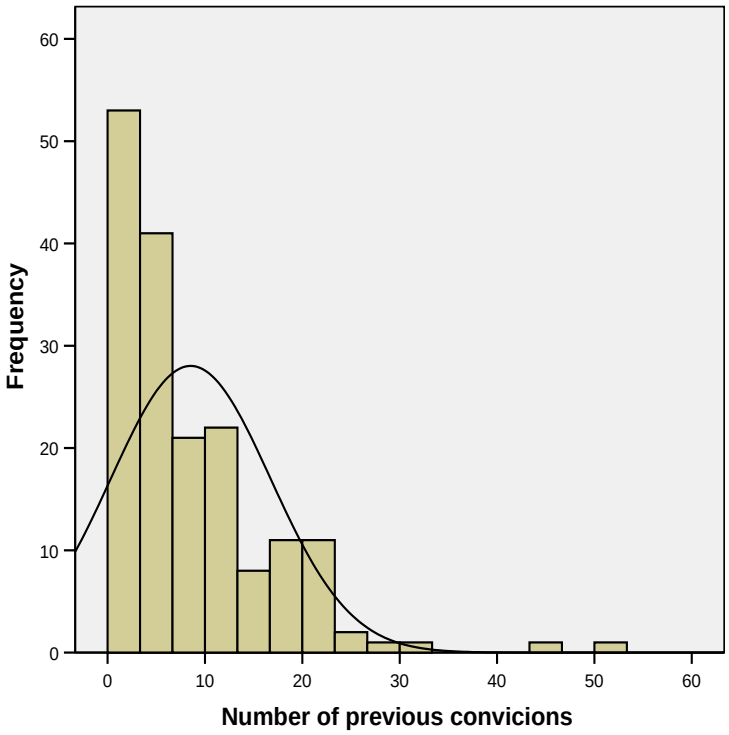
Distribution of age: Comparison group



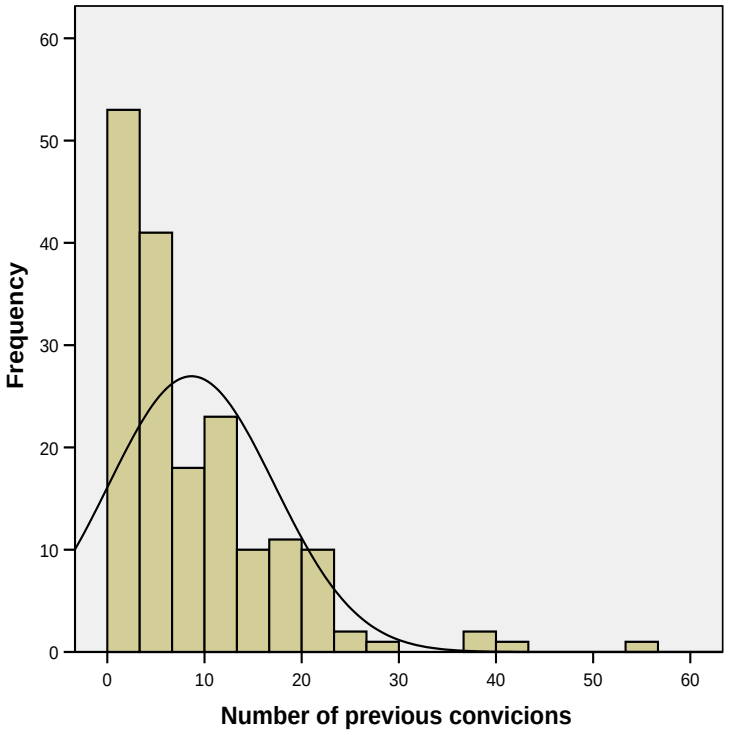
Distribution of age: Experimental group



Distribution of previous convictions: Comparison group



Distribution of previous convictions: Experimental group



Appendix B:

The OGRS2 score is given by the logistical transformation of the following equation

OGRS2 =

$$\frac{\left(100\left(\frac{1}{1+e^{z}}\right)\right)+0.5}{100}$$

Where Z is calculated as below:

$$Z=-3.121+y_1+y_2-0.023\,y_3+0.057*\left(75\sqrt{\left(\frac{y_4}{y_5+5}\right)}\right)+y_6+y_7+y_8+y_9$$

Where the following notation is used:

Notation	Variable
y_1	Age at conviction (years, categorical)
y_2	Gender (1=male 2=female)
y_3	Number of youth custodial sentences
y_4	Total number of court appearances
y_5	Time in years since first conviction
y_6	Age at first conviction (categorical)
y_7	Type of offence (27 categories)
y_8	Any burglary (yes or no)
y_9	Any breach (yes or no)

Offence categories	
Violence	-0.125
Robbery	-0.189
Aggravated burglary	-0.517
Violence against children	-0.491
Public order/riot	-0.135
Firearm	0.094
Sexual	-0.652
Child sex offence	-0.681
Indecent exposure	0.853
Soliciting/prostitution	-1.389
Domestic burglary	0.291
Non-domestic burglary	0.417
Other burglary	0.439
Theft	0.393

Handling	0.321
Fraud and forgery	0.012
Abscond/bail	0.817
TDA and related	0.207
Theft from a car	0.509
Other motoring	0.238
Criminal/malicious damage	0.264
Drugs imp/exp/prod	-0.091
Drugs supply	-0.234
Drgs poss/prem	0.275
Drgs poss/supp	-0.236
Arson	-0.411
other	0.022

Age at conviction	
Less than 14	0.473
14-15	1.290
16-17	0.767
18-20	0.217
21-24	-0.224
25-29	-0.467
30-34	-0.557
35-39	-0.576
40-49	-0.599
50+	-0.323

Age at first conviction	
Less than 14	0.473
14-15	1.290
16-17	0.767
18-20	0.217
21-24	-0.224
25-29	-0.467
30-34	-0.557
35-39	-0.576
40-49	-0.599
50+	-0.323

Any burglary	
Yes	0.187
No	-0.187

Any Breach	
Yes	0.056
No	-0.056

Sex	
Male	0.109
Female	-0.109